

PATENT  
ARAG:003

APPLICATION FOR UNITED STATES LETTERS PATENT

for

LIPOPROTEINS AS NUCLEIC ACID VECTORS

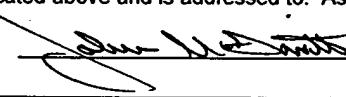
by

Juan G. Guevara, Jr.

Ron C. Hoogeveen

and

Paul J. Moore

EXPRESS MAIL MAILING LABEL	
NUMBER EM423824307US	
DATE OF DEPOSIT May 14, 1998	
I hereby certify that this paper or fee is being deposited with the United States Postal Service "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to: Assistant Commissioner for Patents, Washington D.C. 20231.	
 John McDavitt	

## **BACKGROUND OF THE INVENTION**

The present application is a continuation-in-part of co-pending U.S. Patent Application Serial No. 08/874,807 Entitled "Lipoproteins As Nucleic Acid Vectors" filed June 13, 1997. The entire text of the above-referenced disclosure is specifically  
5 incorporated by reference herein without disclaimer.

### **1. Field of the Invention**

The present invention relates to materials and methods for the *in vivo* transport and delivery of nucleic acids. More particularly, it concerns the use of lipoproteins, 10 including but not limited to, low density lipoproteins ("LDL"), and/or apolipoproteins for the *in vivo* transport of nucleic acids. In addition, the present invention relates to the use of lipoproteins in the early detection of cancer and/or metastatic cancer and/or arteriosclerosis.

### **15 2. Description of Related Art**

The ultimate curative method for any genetic disorder, whether the disorder is inherited or results from a mutation, depends on an effective mode of replacing or augmenting non-functional gene(s). This process is now termed gene or genetic therapy. There are two important aspects to genetic therapy, the gene delivery system/vehicle and 20 the gene control/expression program. Ideally, a replacement gene should become resident in the genome of the target cells/organism and be transferable to subsequent generations of cells and progeny, *i.e.*, the change is incorporated into the germ cells or reproductive cells, the sperm and ovary. Although there have been several significant breakthroughs in this field, this area of biotechnology is still in its early development phase. The first 25 step in any approach to gene replacement is the delivery of the specific gene (nucleic acid) to the cells.

Many techniques have been and are being developed to deliver and express genes in cells and specific tissues in mammals *in vivo*. Several general, non-specific methods

for delivering genes have been reported involving aerosol nucleic acid delivery to cells (Stribling *et al.*, 1992); calcium phosphate precipitation, using a steep change in ionic strength (Wigler *et al.*, 1979); DEAE-dextran (Sompayrac *et al.*, 1981); electroporation, forcing the nucleic acid into the cell by using an electric field or current (Neumann *et al.*, 5 1982); microinjection, physically injecting the nucleic acid into a cell (Benvenisty *et al.*, 1986; Wolff *et al.*, 1990); and polycationic molecules such as polylysine polypeptides (Curiel *et al.*, 1992) and cationic lipids (Lee *et al.*, 1996).

10 Liposomes, vesicles composed of synthetic or non-natural lipids such as long-chain fatty acids, can be used to carry the nucleic acid into the cell cytoplasm non-specifically (Felgner *et al.*, 1987). A recent invention describes the delivery of a self-initiating and self-sustaining gene expression system which contains an RNA polymerase prebound to a DNA molecule using the aforementioned nucleotide delivery systems (U.S. 15 Patent No. 5,591,601).

15 Viral vectors in which specific nucleic acid sequences are incorporated into a neutralized or inactivated virus can use their viral entry mechanism to gain entry to the cell cytoplasm *via* specific cellular receptors to deliver nucleic acids (Schimotohono *et al.*, 1981). The use of specific cellular receptors is apparently a more specific method for 20 delivering genes. In this approach, the nucleic acid is bound either freely, through charge association, or alternatively it is chemically and non-reversibly conjugated to proteins with specific receptor proteins on the membrane of target cells for receptor-mediated uptake (Wu *et al.*, 1988, Wu *et al.*, 1989).

25 Techniques such as calcium phosphate precipitation, electroporation or DEAE-dextran transfection are not suitable for *in vivo* applications. Bombarding cells with nucleic acids under high pressure is a technique which has very limited applications in that it can only be applied topically and only a small number of cells can be targeted. Microinjection of nucleic acids into cells is mainly performed *in vitro* and requires 30 actively dividing cells.

Gene delivery systems that use the viral entry mechanism of recombinant viral vectors have major disadvantages. Systems that utilize replication-defective adenoviral vectors can infect a wide variety of eukaryotic cell types including quiescent somatic cells utilizing the viral entry mechanism. However, adenoviral vector-based delivery systems are only successful in transient gene expression and repeated administration of the viral vector results in a strong immunological response of the host. In addition, the host will experience an adenoviral infection and can experience its symptoms if the recombinant vector undergoes homologous recombination with the wild-type virus strain. Systems that employ recombinant retroviral vectors can be used for stable integration of the gene of interest into the host's genome, but only actively dividing cells can be targeted. In addition, the disadvantages of the adenoviral vector systems also apply to retroviral vector systems (immune response, disease *etc.*).

Positively-charged polycationic molecules such as polylysine peptides which bind non-specifically to the negatively charged nucleic acids have been used to introduce DNA into the chromosome of the recipient cell or organism. Cationic lipid vesicles, liposomes and micelles have been used in aggregates with DNA and viral envelope glycoproteins in non-specific delivery of genes. Liposomes, vesicles composed of synthetic or non-natural lipids, such as long-chain fatty acids, can be used to carry the nucleic acid into the cell cytoplasm non-specifically. In these systems, the liposomes are structured to "best fit" the nucleic acid and insertion into the cell is through non-specific uptake.

The interaction of the liposomal delivery systems discussed above with the nucleic acid to be delivered is non-specific. In addition, prior art techniques are designed to deliver multiple copies of the nucleic acid to the cell cytoplasm. Optimally, however, only one or two copies of a gene should be transfected per cell throughout the organism to replace a defective set of genes only in the specific cells and tissues where it would normally be expressed.

Thus there is a need for a safe and efficient gene delivery system that may be employed in the burgeoning field of gene therapy.

### SUMMARY OF THE PRESENT INVENTION

5 The present invention contemplates a gene delivery system for use in gene therapy. Thus in particular embodiments, the present invention provides a composition comprising an isolated polypeptide comprising at least one LDL or VLDL nucleic acid binding domain; and a nucleic acid comprising an LDL or VLDL binding sequence, wherein the nucleic acid is bound to the polypeptide. In particularly preferred 10 embodiments, the polypeptide comprises an LDL nucleic acid binding domain. In other embodiments, the polypeptide comprises a VLDL nucleic acid binding domain. In particular aspects of the present invention, the nucleic acid comprises an expression region operably linked to a promoter active in eukaryotic cells. In more particular embodiments, the expression region encodes a polypeptide. In other preferred 15 embodiments, the expression region comprises an antisense construct.

In those embodiments in which the expression region encodes a polypeptide, the polypeptide may be selected from the group consisting of  $\alpha$ -globin,  $\beta$ -globin,  $\gamma$ -globin, granulocyte macrophage-colony stimulating factor (GM-CSF), tumor necrosis factor 20 (TNF), IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15,  $\beta$ -interferon,  $\gamma$ -interferon, cytosine deaminase, adenosine deaminase,  $\beta$ -glucuronidase, hypoxanthine guanine phosphoribosyl transferase, galactose-1-phosphate uridyltransferase, glucocerbosidase, glucose-6-phosphatase, thymidine kinase, lysosomal glucosidase, growth hormone, nerve growth factor, insulin, adrenocorticotropic hormone, 25 parathormone, follicle-stimulating hormone, luteinizing hormone, epidermal growth factor, thyroid stimulating hormone of CFTR, EGFR, VEGFR, IL-2 receptor, estrogen receptor, Bax, Bak, Bcl-X<sub>s</sub>, Bik, Bid, Bad, Harakiri, Ad E1B, an ICE-CED3 protease neomycin resistance, luciferase, adenine phosphoribosyl transferase (APRT),

retinoblastoma, insulin, mast cell growth factor, p53, p16, p21, MMAC1, p73, zac1 and BRCA1.

In those embodiments in which the expression region comprises an antisense construct, the antisense construct may be complementary to a segment of an oncogene. 5 In more preferred embodiments, the oncogene may be selected from the group consisting of *ras*, *myc*, *neu*, *raf*, *erb*, *src*, *fms*, *jun*, *trk*, *ret*, *gsp*, *hst*, *bcl* and *abl*.

10 The expression region may be linked to a promoter selected from the group consisting of CMV IE, LTR, SV40 IE, HSV *tk*,  $\beta$ -actin, human globin  $\alpha$ , human globin  $\beta$  and human globin  $\gamma$  promoter. In a defined embodiment, the nucleic acid binding domain is an apoB100 nucleic acid binding domain. In other embodiments, the composition of the present invention may further comprise one or more lipoproteins selected from the group consisting of apoA1, apoA-II, apoA-IV, acat, apoE, apoC-II, apoC-III and apo-D. 15 In particularly preferred embodiment, the apoB100 is selected from the group consisting of human, rat and baboon apoB100.

20 In particular aspects of the invention, the polypeptide comprises at least two nucleic acid binding domains. In particularly preferred embodiments, the nucleic acid binding domain contains a motif selected from the group consisting of a proline pipe helix DNA binding motif, a ISGF3 $\gamma$ -like DNA binding motif, a SREBP-like DNA binding motif, a coiled-coil motif and a nucleotide (ATP)-binding motif. In more defined embodiments, the binding domain may be selected from the group consisting of SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:83, SEQ ID 25 NO:85, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109, SEQ ID NO:110, SEQ ID 30 NO:111, SEQ ID NO:112, NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:144,

SEQ ID NO:145, SEQ ID NO:146, SEQ ID NO:147, SEQ ID NO:148, SEQ ID NO:149, SEQ ID NO:150, SEQ ID NO:151, SEQ ID NO:152, SEQ ID NO:153, SEQ ID NO:154, SEQ ID NO:163, SEQ ID NO:164, SEQ ID NO:165, SEQ ID NO:166 and SEQ ID NO:175.

5

In other embodiments, the polypeptide may further comprise at least one nuclear localization sequence. More particularly, the nuclear localization sequence may be from apoB100. In more preferred embodiments, the nuclear localization sequence may be selected from the group consisting of SEQ ID NO:178, SEQ ID NO: 179, SEQ ID NO: 180, SEQ ID NO: 194, SEQ ID NO: 195, SEQ ID NO: 196, SEQ ID NO: 197, SEQ ID NO: 198, SEQ ID NO: 199, SEQ ID NO: 200, SEQ ID NO: 201, SEQ ID NO: 202, SEQ ID NO: 203, SEQ ID NO: 204, SEQ ID NO: 205, SEQ ID NO: 206, SEQ ID NO: 207, SEQ ID NO: 208, SEQ ID NO: 209, SEQ ID NO: 210.

15

Also contemplated by the present invention is a method for expressing a polypeptide in a human cell comprising the steps of providing a composition comprising (i) an isolated polypeptide comprising at least one LDL or VLDL nucleic acid binding domain and (ii) a nucleic acid comprising an expression cassette comprising a sequence encoding the polypeptide and a promoter active in eukaryotic cells, wherein the coding sequence is operably linked to the promoter, and wherein the nucleic acid sequence is bound to the LDL or VLDL; contacting the composition with the cell under conditions permitting transfer of the composition into the cell; and culturing the cell under conditions permitting the expression of the polypeptide.

25

In particularly preferred embodiments, the polypeptide independently, is a tumor suppressor, a cytokine, an enzyme, a hormone, a receptor, or an inducer of apoptosis. In preferred embodiments, the tumor suppressor may be selected from the group consisting of p53, p16, p21, MMAC1, p73, zac1, BRCA1 and Rb. In preferred embodiments, the cytokine may be selected from the group consisting of IL-2, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, TNF, GMCSF,  $\beta$ -interferon

and  $\gamma$ -interferon. In other preferred embodiments, the enzyme may be selected from the group consisting of cytosine deaminase, adenosine deaminase,  $\beta$ -glucuronidase, hypoxanthine guanine phosphoribosyl transferase, galactose-1-phosphate uridyltransferase, glucocerbrosidase, glucose-6-phosphatase, thymidine kinase and lysosomal glucosidase. In still further preferred embodiments, the hormone may be selected from the group consisting of growth hormone, nerve growth factor, insulin, adrenocorticotropic hormone, parathormone, follicle-stimulating hormone, luteinizing hormone, epidermal growth factor and thyroid stimulating hormone. In defined embodiments, the receptor may be selected from the group consisting of CFTR, EGFR, VEGFR, IL-2 receptor and the estrogen receptor. In other preferred embodiments, the inducer of apoptosis may be selected from the group consisting of Bax, Bak, Bcl-X<sub>s</sub>, Bik, Bid, Bad, Harakiri, Ad E1B and an ICE-CED3 protease.

In particularly preferred embodiments, the nucleic acid binding domain is an apoB100 nucleic acid binding domain. In more preferred embodiments, the apoB100 may be selected from the group consisting of human, rat and baboon low density apoB100. In still further preferred embodiments, the binding region is selected from the group consisting of a proline pipe helix DNA binding motif, a ISGF3 $\gamma$ -like DNA binding motif, a SREBP-like DNA binding motif, a coiled-coil motifs, and a nucleotide (ATP)-binding motif. In particular embodiments, the polypeptide further may comprise at least one nuclear localization sequence. In especially preferred embodiments, the nuclear localization sequence is derived from an apoB100 nuclear localization sequence. In specific embodiments, the polypeptide may be selected from the group consisting of  $\alpha$ -globin,  $\beta$ -globin,  $\gamma$ -globin, neomycin resistance, luciferase, adenine phosphoribosyl transferase (APRT), and mast cell growth factor.

Also provided is a method for providing an expression construct to a human cell comprising providing a composition comprising (i) an isolated polypeptide comprising at least one LDL or VLDL nucleic acid binding domain and (ii) an expression cassette comprising a nucleic acid sequence encoding an expression region and a promoter active

in eukaryotic cells, wherein the expression region is operably linked to the promoter, and wherein the nucleic acid sequence is bound to the LDL or VLDL; contacting the composition with the cell under conditions permitting transfer of the composition into the cell; and culturing the cell under conditions permitting the expression of the expression 5 region.

In particularly preferred embodiments, the expression construct comprises an antisense construct. In more preferred embodiments, the antisense construct is derived from an oncogene. In exemplary embodiments, the oncogene may be selected from the 10 group consisting *ras*, *myc*, *neu*, *raf*, *erb*, *src*, *fms*, *jun*, *trk*, *ret*, *gsp*, *hst*, *bcl* and *abl*. In other embodiments, the expression construct comprises a nucleic acid coding for a gene. In preferred aspects the gene encodes a polypeptide.

In particularly preferred embodiments, the nucleic acid binding domain is an apoB100 nucleic acid binding domain. The apoB100 may be selected from the group 15 consisting of human, rat and baboon low density apoB100. In other preferred embodiments, the DNA binding region is selected from the group consisting of a proline pipe helix DNA binding motif, a ISGF3 $\gamma$ -like DNA binding motif, a SREBP-like DNA binding motif, a coiled-coil motifs, and a nucleotide (ATP)-binding motif.

20

Further the present invention contemplates a method for treating a human disease comprising providing a composition comprising (i) an isolated polypeptide comprising at least one LDL or VLDL nucleic acid binding domain and (ii) an expression cassette comprising a nucleic acid sequence encoding an expression region and a promoter active 25 in eukaryotic cells, wherein the expression region is operably linked to the promoter, and wherein the nucleic acid sequence is bound to the LDL or VLDL; and administering the composition to a human subject having the disease under conditions permitting transfer of the composition into cells of the human subject.

In specific embodiments, the disease may be selected from the group consisting of cancer, diabetes, cystic fibrosis and arteriosclerosis. In preferred embodiments the polypeptide comprises at least two nucleic acid binding regions. In other preferred embodiments the polypeptide comprises at least one nuclear localization sequence. In 5 particularly preferred embodiments, the nucleic acid encodes a gene. In other preferred embodiments, the expression construct comprises an antisense construct.

Another aspects of the present invention describes a pharmaceutical composition comprising an isolated polypeptide comprising at least one LDL or VLDL nucleic acid binding domain; and a nucleic acid comprising an LDL or VLDL binding sequence, 10 wherein the nucleic acid is bound to the polypeptide; the pharmaceutical composition being dispersed in a suitable diluent.

Also contemplated by the present invention is a method of transforming a cell comprising providing a cell; contacting the cell with a composition comprising (i) an isolated polypeptide comprising at least one LDL or VLDL nucleic acid binding domain and (ii) an expression cassette comprising a nucleic acid sequence encoding an expression region and a promoter active in eukaryotic cells, wherein the expression region is operably linked to the promoter, and wherein the nucleic acid sequence is bound to the 20 LDL or VLDL; wherein expression of the expression region is indicative of the transformation.

Yet another aspect of the present invention contemplates a method of transfecting a cell comprising the steps of providing a cell; contacting the cell with a composition comprising (i) an isolated polypeptide comprising at least one LDL or VLDL nucleic acid binding domain and (ii) an expression cassette comprising a nucleic acid sequence encoding an expression region and a promoter active in eukaryotic cells, wherein the expression region is operably linked to the promoter, and wherein the nucleic acid sequence is bound to the LDL or VLDL; wherein expression of the expression region is 25 indicative of the transfection.

Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating preferred 5 embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

#### BRIEF DESCRIPTION OF THE DRAWINGS

10 The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

15 FIG. 1 shows the amino acid sequence of apoB-100.

20 FIG. 2A-FIG. 2G is a homology alignment of SH3-like regions in apo B-100 with known SH3 domains of signal transduction proteins. FIG. 2A-FIG. 2D are the homology alignments and FIG. 2E-FIG. 2G identify the regions of apo B-100 and the proteins aligned.

25 FIG. 3 shows a comparison of SH2-like regions in apo B-100 to known SH3 domains of signal transduction proteins. FIG. 3A and FIG. 3B are the homology alignments, FIG. 3C identifies the proteins and regions aligned.

FIG. 4A-FIG. 4C shows a comparison of the apo B-100 SH 1-like region to SH1 kinase domains of known signal transduction proteins. FIG. 4A and FIG. 4B shows the alignments; FIG. 4C identifies the proteins and regions aligned.

FIG. 5A and FIG. 5B shows the inter-kringle proline-rich regions of Apo[a] compared with the proline rich region of SH3-binding protein (3BP1). FIG. 5A shows the alignment;, FIG. 5B identifies the proteins and regions aligned.

5

FIG. 6 shows an homology alignment of specific regions of apo B-100 and the activation regions located at the amino- and carboxyl- termini of signal transduction proteins.

10

FIG. 7 illustrates the homology of specific regions of apo B-100 with proline pipe helix motifs of Tus.

FIG. 8 shows a homology alignment among one region of the DNA-binding protein ISGF3 $\gamma$  and similar regions in apo B-100.

15

FIG. 9 shows a homology alignment among regions of the DNA-binding protein ISGF3 $\gamma$  and similar regions in apo B-100.

20

FIG. 10A-FIG. 10L. FIG. 10L shows a sequence comparison of the DNA-binding domains of the SREBP1, SREBP2, and ADD1 proteins with similar regions found in apo B-100. FIG. 10B-10L shows a sequence comparison of the DNA-binding domains of SREBP1 with various apolipoproteins.

25

FIG. 11 shows a comparison of the primary structures of known coiled-coil regions of DNA-binding proteins and analogous regions in apo B-100.

FIG. 12 shows a comparison of known ATP-binding loop motifs to similar regions in apo B-100.

FIG. 13 shows a comparison of known nuclear localization signal sequences to similar regions in apo B-100.

5 FIG. 14 shows a comparison of human apo B-100 regions with sequenced regions of apo B-100 from other species.

10 FIG. 15 shows the composition of the LDL gene delivery system of the instant invention LDL containing apo B-100 is depicted along with a DNA sequence containing a promoter, a protein coding region, a 3' untranslated region, and a non-coding region.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

15 The present invention arises from the discovery that regions of apolipoproteins, the protein fraction of lipoprotein particles, are similar in primary structure and amino acid sequence to cellular proteins which are known to bind to DNA. Presently, the only known functions of lipoproteins VLDL, IDL, LDL and HDL are the solubilization and transport of hydrophobic lipids in plasma. The instant invention shows that LDLs, but not other lipoproteins, form a complex with DNA.

20 Herein, synthetic analogues of regions of DNA have been shown to bind to highly purified preparations of human, rat, and baboon LDL but not to other human lipoproteins such as VLDL and HDL, nor to mouse lipoproteins. In fact, the differences observed among the four species tested suggests that human, rat, and baboon lipoproteins behave very similarly in terms of DNA binding preference. Further, purified preparations of human, rat, and baboon LDLs are shown to complex with the promoter region of the 25 human cytomegalovirus. Thus, the present invention demonstrates that human LDL complexes with specific regions of genomic DNA.

Because lipoproteins have specific cell membrane receptors and are actively and specifically internalized by many different cell types in mammals, and because the

inventors show that LDL can bind DNA, these lipoproteins can be used as gene delivery vectors. More specifically, this invention relates to materials and methods for the use of lipoproteins, such as LDL, or, for example, apolipoproteins such as, but not limited to, apoB-100, apoAl, apoE, apoAIV, and apoC, or more specifically still, the DNA binding regions of these lipoproteins, as gene delivery vectors *in vivo*. As explained in greater detail below, the various embodiments of this invention include, but are not limited to, the delivery of nucleic acids to a cell in the form of an LDL-lipoprotein complex, the specific delivery of DNA to the nucleus, and the specific localization of delivered DNA to specific nuclear sites.

10

Plasma levels of DNA increase in a variety of chronic diseases including lupus erythrematosis (Steinman, 1984), viral hepatitis (Neurath *et al.*, 1984), and a variety of cancers (Leon *et al.*, 1977; Shapiro *et al.*, 1983; Stroun *et al.*, 1987; Nawroz *et al.*, 1996; Anker *et al.*, 1997; Chen *et al.*, 1996). It further has been shown that lipoproteins in the blood of non-tumor carrying organisms are not bound to nucleic acids. However, cancer-carrying individuals, and in particular individuals with metastatic cancers, release large amounts of nucleic acids, into the blood. Thus, this invention also relates to the observation that lipoproteins in the blood of cancer patients and especially metastatic cancer patients are bound to nucleic acids, including DNA. Accordingly, this invention also may be used to provide a simple screening test for the presence or absence of cancer, especially metastatic cancer, by isolating a patient's lipoproteins and determining whether the lipoproteins are bound to nucleic acids; the presence of lipoprotein-bound nucleic acid being correlative with the presence of cancer and/or metastatic cancer in the living body. Further embodiments of the present invention relate to the sequence specific detection of DNA bound to lipoproteins in a cancer patient as a method for the identification of specific types of cancer in a living body. These and other aspects of the present invention are discussed in greater detail below.

## 1. LIPOPROTEINS

Lipoproteins appear as micro-pseudomicellar particles in the blood plasma of all mammalian species including humans. Their major function is to transport lipids and other hydrophobic compounds (*i.e.*, fat-soluble vitamins) through the aqueous environment of the blood stream to their specific target cells. The transported lipids can be used as a major substrate for energy metabolism (*i.e.*, triglycerides), structural components for cell membranes (*i.e.*, phospholipids and cholesterol), or as precursors for steroid hormones and bile acids (*i.e.*, cholesterol). Although, lipoproteins vary widely in size and lipid content, they have a common general structure. Lipoprotein particles are believed to be spherical and consist of a hydrophobic core containing nonpolar lipids surrounded by a hydrophilic surface monolayer of polar lipids and proteins, which are called apolipoproteins.

Plasma lipoproteins may be separated into five major classes based on their density, size, and compositional and functional properties: 1) chylomicrons, 2) very low density lipoproteins (VLDL), 3) intermediate lipoproteins (IDL), 4) low density lipoproteins (LDL), and 5) high density lipoproteins (HDL). The different classes of lipoproteins show distinct compositional differences in apolipoprotein content. The specific role of each class of lipoproteins in lipid metabolism is determined by the interaction of these apolipoproteins with specific enzymes and cellular receptors.

### a. ApoB100 Structure and Function

The major protein constituent of LDL is apoB-100. ApoB-100 is one of two known natural ligands for the LDL (apoE/apoB) receptor which is found on the surface of a wide variety of mammalian cell types (Brown and Goldstein, 1986). LDLs are taken up by a process called receptor-mediated endocytosis (Brown and Goldstein, 1986). Hence, lipoproteins may be able to function as naturally-occurring liposomes which contain protein constituents that can bind specifically to nucleic acids and can be internalized by a wide variety of eukaryotic cell types *via* specific receptor mediated processes.

Human apolipoprotein B-100 (apoB-100) is a major apoprotein component of very-low density lipoproteins (VLDL), intermediate density lipoproteins (IDL), low density lipoproteins (LDL), and lipoprotein[a] (Lp[a]). ApoB-100 is synthesized and incorporated into VLDL and Lp[a] by the liver. Human LDL can be described as a spherical particle composed of a hydrophobic core of cholesterol esters and triglycerides encapsulated by an amphipathic monolayer of phospholipids, glycolipids and cholesterol in which the apoB-100 is partially imbedded (Myant, 1990). In addition to one molecule of apoB-100, LDL is known to contain varying numbers of apo C-I, apo C-II, apo C-III, apo E, and apo D (Blanco-Vaca *et al.*, 1992; Connelly *et al.*, 1993; Blanco-Vaca *et al.*, 10 1994).

The primary structure of apoB-100, SEQ ID NO:1 (FIG. 1) has been determined by amino acid sequence analysis (Yang *et al.*, 1986; Yang *et al.*, 1989) and inferred from its cDNA sequence (Yang *et al.*, 1986; Yang *et al.*, 1989; Knott *et al.*, 1986). There 15 appear to be several different isoforms of apo B-100. The isoform shown in FIG. 1 is the isoform used for all of the alignments in the specification. Homologous regions in the other isoforms, however, would align similarly.

The apparent molecular weight of apoB-100 is 512 kDa based on its amino acid 20 composition of 4536 residues. The apoprotein contains 25 Cys residues (Coleman *et al.*, 1990; Yang, 1990), at least 16 of which form intramolecular disulfide bonds, with the remaining cysteines present as free sulphydryls, as additional (unassigned) intramolecular disulfides, or as intermolecular disulfide linkages to other apolipoproteins (Blanco-Vaca *et al.*, 1992; Connelly *et al.*, 1993). Several important functional regions on apoB-100 25 that have been identified include heparin-binding sites (Cardin *et al.*, 1987; Weisgraber and Rall, 1987), glycosylation sites (Knott *et al.*, 1986; Innerarity *et al.*, 1986), and the LDL receptor-binding region (Blanco-Vaca *et al.*, 1992, Knott *et al.*, 1986, Milne *et al.*, 1989).

ApoB-100, and apolipoprotein E (apoE), apolipoproteins present in the low-density lipoprotein group, function as ligands for the high-affinity receptor-mediated removal of certain lipoproteins from plasma by the liver and delivery of cholesterol and cholesterol esters to a variety of target tissues (Myant, 1990; Innerarity *et al.*, 1986; 5 Brown and Goldstein, 1986; Mahley, 1988). A general mechanism for the receptor mediated uptake of LDL is well-established (Myant, 1990; Innerarity *et al.*, 1986; Brown and Goldstein, 1986; Mahley, 1988), and the role of the apoB-100 molecule in this mechanism also is well defined.

10 Specific binding of low density lipoproteins to their mammalian cell receptors depends on the presence and conformation of the apoB-100 ligands (Kinoshita *et al.*, 1990). Several reports have shown that the interaction of apoB-100-lipoproteins with the up-regulated, high affinity LDL (apoB/apoE) receptor is modulated by the lipid composition of the particle (Teng *et al.*, 1985; Marcel *et al.*, 1988), by other apoproteins 15 such as apo[a] in Lp[a] (Kostner and Grillhofer, 1991; Young *et al.*, 1986) and apoE in  $\beta$ -VLDL (Innerarity *et al.*, 1986; Mahley, 1988), and by monoclonal antibodies to specific regions of the apoB-100 molecule (Innerarity *et al.*, 1986; Young *et al.*, 1986).

20 In searching the apoB-100 sequence for regions of sequence similarity to other proteins, nucleic acid binding regions (deoxyribonucleic acids, DNA and ribonucleic acids, RNA), nucleotide-binding regions, and nuclear-localization regions in the amino acid sequence of apoB-100 and apoE, have been identified. The present invention demonstrates that highly purified preparations of human, rat, and baboon LDL bind 25 specifically to pure preparations of human genomic DNA. These properties impart to the lipoproteins the capacity to serve as delivery vehicles for genetic material.

30 Lipoprotein particles carry a variety of vitamins and steroid compounds in their pseudo-micelle lipid core which may function in the control of gene expression. These attributes impart to the lipoproteins a virus-like character as well as capacity. While the inventors do not wish to be bound by any particular theory, the many control elements

and signal motifs in the primary structure of the apolipoproteins are suggestive of the ability of these proteins to transport nucleic acids, enter the cell, participate in signal transduction, enter the nuclear space, initiate incorporation of nucleic acid materials into the resident genome, and cause its subsequent expression. As used herein, the term  
5 "primary structure" refers to the amino acid sequence of the protein. The capacity of purified LDL to bind to human genomic DNA, along with apoB-100's homology to SH1, SH2, and SH3 signal transducer domains supports this hypothesis. These properties of apoB100, and methods of exploiting these properties, are discussed in further detail below.

10

## **2. NUCLEIC ACID BINDING REGIONS**

The inventors have found that apo B-100 is also involved in DNA binding. DNA is the genetic blueprint that contains the information necessary for cell growth, differentiation, proliferation, and cellular response to environmental factors. The  
15 phenotypic differences between various cell types in higher eukaryotes are mainly due to differences in cellular gene expression.

The regulation of gene expression is predominantly controlled at the stage of initiation of transcription and is mediated by proteins which recognize specific DNA  
20 sequences. In order to recognize and bind to a specific DNA sequence a protein utilizes a structural motif. Over the past 15 years, several structural DNA binding motifs have been identified including as zinc fingers, helix-turn-helix, basic helix-loop-helix, KH RNA-binding motifs and leucine zippers and proline pipe helices. The inventors report here the identification of regions in apo B-100 with homology to various DNA binding motifs  
25 including: 1) Proline pipe helix DNA binding motifs, 2) ISGF3 $\gamma$ -like DNA binding motifs, 3) SREBP-like DNA binding motifs, 4) coiled-coil motifs, and 5) nucleotide (ATP)-binding motifs.

### **a. Nucleotide and ATP Binding Motifs**

The inventors discovered that that there is a certain degree of homology between regions of apo B-100 and known ATP binding motifs found in other proteins including those involved in signal transduction and transcriptional-ribonucleotide synthesis (t-RNA synthetases. Typically, these proteins contain sites which interact with different regions 5 of the nucleotide, *i.e.*, negatively charged phosphate regions, the ribose (carbohydrate) hydroxyl groups, and the base. A second site binds to the substrate ligand such as any amino acid in the case of t-RNA synthetases and tyrosine, serine and threonine residues in the phosphorylation of proteins.

10 Examination of the apoB-100 primary structure reveals several regions which are similar in sequence to the known nucleotide and ATP binding motifs and are suggestive of a similar function. For example, ATP-binding sites are known to contain an essential ATP-binding lysine residue. In *lyn*, the site is T<sub>269</sub>KVAVTLKPG (SEQ ID NO:54) and in *lyk*, it is D<sub>386</sub>KVAIKTIREG (SEQ ID NO:55). A similar region can be found in apoB-15 100, DLNAVANKIAD (SEQ ID NO:56). The similarity of this region in apo B-100 with the ATP-binding sites on known tyrosine-kinases suggests that apo B-100 can bind to the nucleic acid, ATP.

20 A single ATP-binding region occurs between residues 3800 and 3840 which is located in the kinase domain of apoB-100. The sequence of this region with known ATP-binding regions of kinases is shown in FIG. 12. FIG. 12 shows a comparison of known 25 ATP-binding loop motifs to similar regions in apo B-100. Bold letters indicate conserved amino acids, critical amino acids (H and K) are indicated by the #, "\*" indicates conserved amino acids, "-" indicates gaps introduced in the sequence in order to align the proteins, and identical amino acids between the sequences in "C" are listed below the alignment. Sequence identification numbers are listed in the right margin. The critical lysine residue is retained and the degree of similarity suggests a like function.

30 The ATP-binding motifs typical of t-RNA synthetases are characterized by the signature sequence HIGH (histidine, isoleucine, glycine histidine) SEQ ID NO:177, and a

second motif which contains a critical lysine residue. These motifs are located within 300 residues and occur as proximal loops on the surface of the protein molecule. Several analogues of this signature sequence occur in the apoB-100 sequence (see FIG. 7 and FIG. 12). An extended comparison of apoB-100 regions which contain the HIGH  
5 signature sequence is made with the tyrosyl-tRNA synthetase sequence shown in FIG. 12.

**b. Proline Pipe Helix Structures**

The proline pipe helix is usually present in proteins that contain proline every fifth position (Myant, 1990) in the amino acid sequence that is at least 20 residues long  
10 (PXXXXP)<sub>n</sub> (SEQ ID NO:75) where n>4. In the proline pipe helix, 5.56 residues are required to make one complete left handed helical turn. The proline pipe helix is stabilized by a hydrogen bonding network between the C=O groups of residues in positions i+ 1, i+2, i+3 (where i is a proline or sometimes non-proline residue) with the NH groups in positions i+2, i+3, i+4, respectively, of the following turn (Blanco-Vaca *et al.*, 1992). The unusually large turn of the helix results in the formation of a channel running along the helix that is about 6Å in average diameter (Myant, 1990) and large enough to accommodate water (Blanco-Vaca *et al.*, 1992) and possibly other molecules.  
15

One function of the proline pipe helix is DNA binding. For example, the proline  
20 pipe helix in *Tus* is involved in tight binding to highly specific 22-23 base pair DNA known as *Ter* sites (Connelly *et al.*, 1993; Blanco-Vaca *et al.*, 1994). Because of its large diameter compared to the  $\alpha$ -helix, the proline pipe helix spans the entire width of the major groove (Blanco-Vaca *et al.*, 1992) and results in a tight and highly specific fit. This tight fit also results in a high correspondence between the positively charged amino  
25 acid residues of the proline pipe helix and the negatively charged phosphate groups of DNA (Blanco-Vaca *et al.*, 1992). The occurrence of the proline pipe-DNA interactions in nature might be more widespread than presently thought and this interaction might play a very important biological function.

Careful examination and analysis of the apoB-100 amino acid sequence shows that the 40-residue proline-rich segment P2682-I2719, or a portion of this segment, assumes a proline pipe helical conformation (see FIG. 7), PDFRLPEIAIPEFIPIITLNLDQVQVPLHIEFQLPHISH (SEQ ID NO:76). Because 5 the unique features of the proline pipe helix make it suitable for tight and highly specific DNA binding, this segment or motif in apoB-100 constitutes one of the DNA binding sites.

10 The functional implications of DNA binding by apoB-100 include, but are not limited to: 1) binding of DNA such as, for example, microsatellite DNA (Connelly *et al.*, 1993; Blanco-Vaca *et al.*, 1994) to apoB-100 or its fragment(s) for DNA transport from the cytoplasm to the nucleus; (2) binding of apoB-100 or its fragment(s) to the nuclear DNA to regulate transcription or effect other functions; or (3) binding of DNA to apoB-100 or its fragment(s) to transport DNA from the nucleus to the cytoplasm. Other 15 functions as a consequence of apoB-100 DNA binding through the apoB-100 proline pipe helix are not precluded. Therefore, the proline pipe region of apoB-100 constitutes an important target for structure-based drug design and delivery systems.

20 **c. ISGF3 $\gamma$ -like DNA binding motifs**

ISGF3 is a multimeric transcription factor involved in the regulation of transcription of a large set of genes. This factor dissociated into two protein components termed ISGF3 $\gamma$  and ISGF3 $\alpha$ . ISGF3 $\gamma$  is a 48 kDa protein that binds DNA recognizing the IFN-stimulated response element. ISGF3 $\alpha$  does not bind DNA. Regions in apoB-100 have been found to be homologous to the DNA-binding domain of ISGF3 $\gamma$  (FIG. 8 25 and FIG. 9).

FIG. 8 shows a homology alignment among one region of the DNA-binding protein ISGF3 $\gamma$  and similar regions in apo B-100. Basic amino acids are indicated in bold and \* indicates conserved amino acids between the two regions and V indicates

conserved amino acids that have switched positions between the two sequences aligned. Sequence identification numbers are identified in the legend to the figure.

5 FIG. 9 shows a homology alignment among regions of the DNA-binding protein ISGF3 $\gamma$  and similar regions in apo B-100. Basic amino acids are indicated in bold, "-" indicates gaps introduced in the sequence in order to align the two proteins. Sequence identification numbers are identified in the right margin.

This indicates apoB-100 can bind specific DNA sequences in a manner similar to ISGF3 $\gamma$ .

10

#### d. SREBP-Like DNA Binding Motifs

Another region within apoB-100 shows striking resemblance to the DNA binding domains of previously identified sterol regulatory element binding proteins (SREBP's; FIG. 10A and FIG. 10B). A sequence comparison of the DNA-binding domains of the 15 SREBP1, SREBP2, ADD1 proteins with similar regions found in apo B-100 are shown in FIG. 10A where basic amino acids are indicated in bold, "\*" indicates conserved amino acids, "-" indicates gaps introduced in the sequence in order to align the two proteins, and identical amino acids between the two sequence are listed below the alignment. FIG. 10B shows a sequence comparison of the DNA-binding domains of SREBP1 with various 20 apolipoproteins where basic amino acids are indicated in bold, "\*" indicates conserved amino acids, "-" indicates gaps introduced in the sequence in order to align the two proteins, V indicates conserved amino acids that have switched positions between the two sequences aligned, and identical amino acids between the two sequences are listed below the alignment. Sequence identification numbers are indicated in the legend to the figure. 25 The full line of "\*\*\*\*\*" separates the different sequence alignments.

30 SREBP's are members of the basic helix-loop-helix-leucine zipper (bH-L-H-Zip) family of transcription factors and play a major role in the transcriptional regulation of a number of genes involved in cholesterol homeostasis as well as lipid biosynthesis. SREBP's contain 3 segments: 1) an NH<sub>2</sub> terminal bH-L-H-Zip DNA binding domain

including an acidic transcription activating domain; 2) a middle segment containing two membrane spanning domains; and 3) a COOH terminal segment. In order for SREBP's to become functionally active transcription factors, their NH<sub>2</sub> terminal domain containing the bH-L-H-Zip region needs to be released from the endoplasmic reticulum or nuclear envelope. This process is mediated by a sterol-regulated protease. That apo B-100, like the SREBP's, binds DNA.

5 **e. Coiled-coil Motif (Leucine Zipper)**

10 The coiled-coil motif (Myant, 1990), sometimes referred to as the leucine zipper (Blanco-Vaca *et al.*, 1992), is characterized by two  $\alpha$ -helical chains that wrap around each other to form a left-handed supercoil. The amino acid sequence of coiled-coil forming proteins is characterized by the presence of heptad repeats, that is, three or more repeats of a seven-residue sequence where every third and every fourth position in the heptad is occupied by a hydrophobic residue (Blanco-Vaca *et al.*, 1992; Connelly *et al.*, 1993; Blanco-Vaca *et al.*, 1994). The two  $\alpha$ -helical chains that form the coiled-coil can align either in parallel or anti-parallel orientation and their stabilities are dependent on the presence of strategically located hydrophobic and electrostatic interactions (Yang *et al.*, 1986; Yang *et al.*, 1989; Knott *et al.*, 1986; Coleman *et al.*, 1990; Yang, 1990; Cardin *et al.*, 1987; Weisgraber and Rall, 1987; Innerarity *et al.*, 1986; Milne *et al.*, 1989; Brown 15 and Goldstein, 1986). The most attractive feature of the coiled-coil is that highly specific interactions can be tailored by redesigning this relatively simple motif.

20

25 The coiled-coil motif occurs widely in native proteins (Lupas *et al.*, 1991; Cohen and Parry, 1986). It plays structural and functional roles in fibrous proteins such as keratin, myosin, elastin, fibrinogen, tropomyosin, *etc.* The coiled-coil motif also serves as the dimerization domain for a number of transcription factors such as GCN4 (O'Shea *et al.*, 1991; Ellenberger *et al.*, 1992), GAL4 (Kraulis *et al.*, 1992; Baleja and Sykes, 1991; Marmorstein *et al.*, 1992), c-Fos-c-Jun (Glover and Harrison, 1995), where only the dimeric form binds to DNA and is active. It is found in globular proteins, such as 30 tRNA synthetase (Cusack *et al.*, 1990; Biou *et al.*, 1994), and serves as anchors into the

tRNA. Naturally occurring coiled-coils can also be found as three-stranded (Bullough *et al.*, 1994a; Bullough *et al.*, 1994b) or four-stranded (Banner *et al.*, 1987) structures.

Sequence alignment analysis of apoB-100 predicts that there are at least eight coiled-coil structures of varying lengths in different regions of its amino acid sequence (FIG. 11). FIG. 11 shows a comparison of the primary structures of known coiled-coil regions of DNA-binding proteins and analogous regions in apo B-100. Bold letters indicate conserved amino acids. Sequence identification numbers are listed in the right margin.

10

While the inventors do not wish to be bound by any particular theory, it is likely that these coiled-coil domains play very important structural and functional roles that, in turn, are vital to the function of LDL. For example, the coiled-coil motif can serve as dimerization or multimerization sites that may be important in LDL solubilization or aggregation. The coiled-coil motif can also bind DNA, RNA or nucleotides and, therefore, plays a very important role in the regulation and energetics of protein synthesis. The coiled-coil motif can also serve as a template for transport of molecules within and between the cytoplasm and the nucleus. In addition, the coiled-coil motif can also serve as a (temporary) reservoir of ligands that may be important in the regulation of the metabolic pathways. This list is by no means exhaustive, but demonstrates the biological importance of the coiled-coil motif in apoB-100.

The discovery of the coiled-coil motif in apoB-100 and the important biological implications of its presence, apoB-100 by itself or as part of LDL, constitutes an important target for structure-based drug design, delivery, and diagnostic systems. Coiled-coil forming sequence in apoB-100 (as indicated in FIG. 11) can be used to design, study and manufacture coiled-coil based peptide or protein delivery systems for drugs, radioisotopes, oligonucleotides, genes, antigens, antibodies, epitopes for vaccines, sugars, carbohydrate analogs and other ligands to specific targets in cells, tissues and organs. Either single strand or multiple strands of the apoB-100 coiled-coil forming

peptide sequences that can be used as components of or attached to the aforementioned ligands either by covalent or non-covalent methods.

Coiled-coil forming sequences in apoB-100 (FIG. 11), or fragments, analogs, or 5 modifications therefore can be used as site-specific targets for the delivery of drugs, radioisotopes, oligonucleotides, genes, antigens, antibodies, epitopes for vaccines, sugars, carbohydrate analogs and other ligands. Site-specific targeting includes the use of coiled-coils, coiled-coil forming peptides, or any functional group that binds to the aforementioned coiled-coils sequences in apoB-100.

10

### 3. NUCLEAR LOCALIZATION SIGNALS

In addition to homology with DNA binding proteins, apoB-100 contains several 15 regions that are homologous to known nuclear localization signals (FIG. 13). These signals include the NLS from human p53, Abl, and apoJ. FIG. 13 shows a comparison of known nuclear localization signal sequences to similar regions in apo B-100.

20

The bipartite nuclear localization signal contains two essential elements comprised of basic amino acids, H (histidine), R (Arginine), and K (Lysine) which are required for nuclear targeting. The signal motifs starts with two basic amino acids which are then followed by a ten to thirty amino acid spacer and a basic duster of five amino acids three of which must be basic. Approximately 50% of the known nuclear proteins listed in the protein databases have this motif, while less than 5% of non-nuclear proteins have it. FIG. 13A shows sequences in apoB-100 with the perfect 10 amino acid spacer between the bipartite nuclear localization sequence element.

25

There is no strict requirement for the spacer length other than perhaps flexibility in the amino acids, *i.e.*, the dihedral angles. Indeed, there are basic amino acid clusters in the apo B-100 molecule that are separated by longer spacers and are nevertheless potential DNA-binding regions. FIG. 13B shows sequences in apoB-100 with more or 30 less than 10 amino acids in the spacer region between the bipartite nuclear localization

sequence element, and FIG. 13C shows sequences in apoB-100 with more or less than 10 amino acids in the spacer region between an imperfect "bipartite" nuclear localization sequence element.

5           Thus, these regions in apoB-100 are NLS sequences capable of directing DNA to the nucleus of a cell. Apolipoproteins present on human LDL can bind to DNA through the DNA binding motifs identified herein. The functional bH-L-H-Zip domain within apoB-100 can enter the nucleus, following proteolytic release and/or aided by the nuclear localization signal domains present within the apolipoproteins, and regulate transcription  
10           of the target genes.

15           In addition, apo B-100 appears to be conserved across species. FIG. 14 shows various regions of human apo B-100 aligned with the sequenced fragments of the apo B-100 from pig, rat, hamster, mouse, chicken and rabbit. Bold and underlined letters indicate positively charged, basic amino acids, and "—" indicates gaps introduced in the sequence in order to align the proteins;

#### 4. HOMOLOGY TO SIGNAL TRANSDUCING PROTEINS

20           The inventors have found that in addition to homology with nuclear localization signals and DNA binding proteins, apoB-100 molecule has regions of sequence similarity to known motifs in a variety of signal transduction molecules. For example, regions of apo B-100 are homologous to src homology 3 (SH3) (FIG. 2), src homology 2 (SH2) (FIG. 3) and src homology 1 (SH1) (FIG. 4) kinase domains that are common to protein tyrosine kinases of the signal transduction system (Koch *et al.*, 1991; Pawson, 1992; Schlessinger, 1994; Margolis, 1992; Waksman *et al.*, 1993; Carpenter, 1992; Ugi *et al.*, 1994; Lowenstein *et al.*, 1992; Guevara, Jr. *et al.*, 1994), as well as activation regions located at the amino-and carboxyl- termini of signal transduction proteins (FIG. 6).

30           FIG. 2 is a homology alignment of SH3-like regions in apo B-100 with known SH3 domains of signal transduction proteins, where "\*" indicates conserved amino acids,

5       "-" indicates gaps introduced in the sequence in order to align the two proteins, identical amino acids between the two sequences are listed below the alignment, and percent similarity is indicated in the right margin. This alignment is followed by a table identifying the regions of apoB-100 and the various proteins aligned to these regions along with their respective sequence identification numbers.

10       FIG. 3 shows a comparison of SH2-1like regions in apo B-100 to known SH3 domains of signal transduction proteins, where structurally important motifs are indicated by double underline, basic amino acids are indicated in bold, "\*" indicates conserved amino acids, "-" indicates gaps introduced in the sequence in order to align the two proteins, identical amino acids between the two sequences are listed below the alignment, and percent similarity is indicated in the right margin. The alignment is followed by a table identifying the reference proteins and regions of apoB-100 in the alignment along with their sequence identification numbers.

15

20       FIG. 4 shows a comparison of the apo B-100 SH1-like region to SH1 kinase domains of known signal transduction proteins where basic amino acids are indicated in bold, "\*" indicates conserved amino acids, "-" indicates gaps introduced in the sequence in order to align the two proteins, and identical amino acids between the two sequences are listed above the alignment. The alignment is followed by a table identifying the reference proteins and the region of apoB-100 used for the alignment along with their respective sequence identification numbers.

25       FIG. 6 shows a homolog alignment of specific regions of apo B-100 and the activation regions located at the amino- and carboxyl- termini of signal transduction proteins where "\*" indicates conserved amino acids, "-" indicates gaps introduced in the sequence in order to align the two proteins, and identical amino acids between the two sequences are listed above the alignment. Numbers in parenthesis indicate amino acid residues shown in the alignment and sequence identification numbers are listed in the right margin.

Discovery of these motifs in the apoB-100 sequences was based on a series of reports (Ye *et al.*, 1988; Trieu and McConathy, 1990; Trieu *et al.*, 1991) which showed that free proline inhibited binding of recombinant apo[a] to both Lp[a] and LDL. These 5 results implied that proline within the apoB-100 sequence interacted with the kringle binding pocket. Molecular modeling was used to determine if proline is a ligand for the different apo[a] kringle types (Guevara, Jr. *et al.*, 1993). These studies concluded that although free proline can be accommodated by the ligand binding site of several apo[a] kringle types, proline located within a polypeptide chain probably does not fit into any of 10 the ligand binding sites of apo[a] kringles. As an alternative possibility, proline might bind at an allosteric site on the kringle structure (Guevara, Jr. *et al.*, 1993), and thereby alter the ligand binding site of the kringle. A second possibility is that apo[a] kringles are not involved at all, but rather that the proline/threonine-rich inter-kringle regions (IKR's) 15 associate with specific sites on apoB-100, and thereby enable recombinant apo[a] to bind to Lp[a] and LDL.

**a. The SH3 Domain**

The interkringle regions of Apo [a] have homology to 3BP1 (FIG. 5). FIG. 5 shows the inter-kringle proline-rich regions of Apo[a] compared with the proline rich 20 region of SH3-binding protein (3BP1) where the conserved prolines are indicated in bold and "-" indicates gaps introduced in the sequences in order to align the two proteins. Following the alignments is a table identifying the inter-kringle proline-rich regions of Apo[a] and the proline-rich region of 3BP1 used for the alignment along with their respective sequence identification numbers.

25

Apo[a] is a hydrophilic, glycosylated apoprotein that is disulfide-linked to apo B-100 in the Lipoprotein[a] particle. The proline-rich hinge between kringle structures of the apo[a] are suggestive a of role in signaling. Cicchetti *et al.* (1992) and Ren *et al.* (1993) described a ten amino acid, proline-rich segment of the 3BP-1 protein which binds 30 to an SH3 domain in Abl, a non-receptor protein tyrosine kinase involved in signal

transduction. The proline-rich IKR's in apo[a] (McLean *et al.*, 1987; Guevara, Jr. *et al.*, 1992), like those in 3BP-1, contain the sequence PXP (SEQ ID NO:2) which is important for the interaction of these motifs with their corresponding SH3 domains.

5 Proline-rich binding proteins (BP's), SH3, and SH2 domains are regulatory domains in signaling proteins which mediate enzymatic activity, participate in intracellular protein-protein interactions, and bind to activated receptor protein-tyrosine kinases (Koch *et al.*, 1991; Pawson, 1992; Schlessinger, 1994; Margolis, 1992; Waksman *et al.*, 1993; Carpenter, 1992; Ugi *et al.*, 1994; Lowenstein *et al.*, 1992; Guevara, Jr. *et al.*, 1994; Pleiman *et al.*, 1994). The sequence similarities noted between apo[a] IKR's and the proline-rich segment of 3BP-1 suggest a similar function for these regions of the apo[d] in non-covalent interactions between apo[a] and apoB-100, *i.e.*, binding of a proline-rich region in apo[a] to an SHB-like region in apoB-100.

10

15 In apoB-100, at least 13 regions share high sequence similarities with SH3 domains. SH3 domains are found in several signal transduction proteins such as phosphatidylinositol-3' kinase (PI3K) and the non-receptor tyrosine kinase Abl (see FIG. 1 and FIG. 4). This suggests that apo B-100 may have signal transduction properties.

20 **b. The SH2 Domain**

Many signal transduction proteins and other proteins such as tyrosine phosphatases and tensin also contain SH2 domains (Koch *et al.*, 1991; Pawson, 1992; Schlessinger, 1994; Lowenstein *et al.*, 1992), often flanked by SH3 domains. SH2 domains are typically comprised of about 100 amino acids. In the signaling process, SH2 domains bind to specific phosphotyrosine motifs of target proteins (Songyang *et al.*, 1993; Escobedo *et al.*, 1991). The apoB-100 sequence was examined for presence of SH2-like regions and numerous regions in the apoB-100 sequences were found to share some commonalities with SH2 domains of signaling proteins (FIG. 3). This suggests that apoB-100 may interact with phosphorylated proteins through SH2-like regions.

25

### c. The SH1 Domain

Typically, signal transduction proteins also contain a kinase domain or src homology domain 1 (SH1) which is located in the carboxyl region of the protein and is comprised of about 300 amino acids (Rudd *et al.*, 1993). SH1 domains are highly homologous. Regions of apo B-100 have been found that share homology with SH1 domains (FIG. 4). In addition, apo B-100 shares homolog with the catalytic loop or active site motif in these signaling proteins. For example, the active site motif of *lyn* (EC 2.7.1.1 12) is R<sub>359</sub>KNYIHRDLRAAN (SEQ ID NO:52); a sequence that is highly conserved. An analogous region is found in apoB-100, K<sub>3919</sub>GTLAHRDFSAE (SEQ ID NO:53).

Furthermore, apo B-100 shares amino acid sequence homolog with the activation regions located at the amino- and carboxyl- termini of signal transduction proteins (FIG. 15 6). Protein Kinase C and c-AMP-dependent kinase control sites are present at the amino-terminus of signal transduction proteins. Tyrosine kinase control sites are located in the carboxyl-terminus of these proteins. Typically, there is little sequence homology, at the amino-termini, but high homology is common at the carboxyl-termini of signaling protein kinases.

20

Regions of homology, within apo B-100 having sequence similarity to SH3, SH2 and SH1 domains and other cell signaling proteins, all point to the possibility that apo B-100 is involved in intracellular signaling.

## 25 5. PROTEIN EXPRESSION

As described above, the inventors have discovered that a particular region of the apoB-100 molecule is similar in sequence to the Steroid Regulatory Element Binding Proteins, SREBP1 and 2 and ADD1. Other regions of the apoB-100 molecule are similar to specific regions in other known DNA binding proteins including, but not limited to ISGF3 $\gamma$ , coiled-coil regions of GCN4 and hMLKI, and the proline-pipe sequences of Tus.

Further, the inventors found that the amino acid sequence of apolipoproteins, such as apoB-100 have regions involved with nucleotide binding and nuclear localization. For example, apolipoproteins such as apoB-100 show homology to the SH1 kinase domains of protein tyrosine kinases and the HIGH and KMSK motif plus critical lysine of tRNA synthetases both known to bind ATP as well as to the basic helix-loop-helix motif of sterol regulatory element binding proteins (SREBPs) known to localize to the nucleus where they are involved in the regulation of transcription.

5  
10  
a. **Expression of apoB100**

In certain embodiments of the present invention, it will be necessary to obtain apoB100 or lipoproteins containing apoB100 for use as DNA binding compositions. In particular embodiments as described herein below, such apoB100 may be obtained from the lipoprotein fraction of primate serum. As an alternative to purifying apoB100 from LDL fraction of serum, it is possible to generate pure fractions of apoB-100 by recombinant expression of the apoB100 gene. The apoB100 gene can be inserted into an appropriate expression system. The gene can be expressed in any number of different recombinant DNA expression systems to generate large amounts of the polypeptide product, which can then be purified and used as a DNA binding composition as described herein.

20  
25  
In one embodiment, specific amino acid sequence domains of an apoB100 polypeptide having for example, the sequence of SEQ ID NO:1, can be prepared. These may, for instance, be minor sequence variants of a polypeptide that arise due to natural variation within the population or they may be homologues found in other species. They also may be sequences that do not occur naturally but that are sufficiently similar that they function similarly and/or elicit an immune response that cross-reacts with natural forms of the polypeptide.

30  
The nucleotide binding, nuclear localization and signal transduction domains of the apoB100 molecule are discussed in detail herein below. Recombinant technologies, well known to those of skill in the art, may be used to produce recombinant apoB100

with one or more of these domains having sequences that optimize the DNA binding and/or nuclear localization capacities of the molecule. Furthermore, in certain instances it may be necessary to "customize" such domains in order to increase binding to a particular DNA sequence whilst decreasing the binding to other sequences. Alternatively, it may be 5 preferable to alter a particular apoB100 polypeptide, in order to decrease its binding affinity for a particular molecule. Accordingly, sequence variants of these domains can be prepared by standard methods of site-directed mutagenesis such as those described below in the following section.

10 Amino acid sequence variants of an apoB100 polypeptide, or particular domains therein can be substitutional, insertional or deletion variants. Deletion variants lack one or more residues of the native protein which are not essential for function or immunogenic activity.

15 Substitutional variants typically contain the exchange of one amino acid for another at one or more sites within the protein, and may be designed to modulate one or more properties of the polypeptide such as stability against proteolytic cleavage. Substitutions preferably are conservative, that is, one amino acid is replaced with one of similar shape and charge. Conservative substitutions are well known in the art and include, for example, the 20 changes of: alanine to serine; arginine to lysine; asparagine to glutamine or histidine; aspartate to glutamate; cysteine to serine; glutamine to asparagine; glutamate to aspartate; glycine to proline; histidine to asparagine or glutamine; isoleucine to leucine or valine; leucine to valine or isoleucine; lysine to arginine; methionine to leucine or isoleucine; 25 phenylalanine to tyrosine, leucine or methionine; serine to threonine; threonine to serine; tryptophan to tyrosine; tyrosine to tryptophan or phenylalanine; and valine to isoleucine or leucine.

30 Insertional variants include fusion proteins such as those used to allow rapid purification of the polypeptide and also can include hybrid proteins containing sequences from other proteins and polypeptides which are homologues of the polypeptide. For

example, an insertional variant could include portions of the amino acid sequence of the polypeptide from one species, together with portions of the homologous polypeptide from another species. Other insertional variants can include those in which additional amino acids are introduced within the coding sequence of the polypeptide. These typically are 5 smaller insertions than the fusion proteins described above and are introduced, for example, into a protease cleavage site. Alternatively, insertional variants of the present invention may be created in which one or more DNA binding domains and nuclear localization domain have been added to a native apoB100 molecule to alter particular characteristics of the molecule.

10

In one embodiment, major antigenic determinants of the polypeptide are identified by an empirical approach in which portions of the gene encoding the polypeptide are expressed in a recombinant host, and the resulting proteins tested for their ability to elicit an immune response. For example, PCR can be used to prepare a range of cDNAs encoding 15 peptides lacking successively longer fragments of the C-terminus of the protein. The immunoprotective activity of each of these peptides then identifies those fragments or domains of the polypeptide that are essential for this activity. Further experiments in which only a small number of amino acids are removed at each iteration then allows the location of the antigenic determinants of the polypeptide.

20

Another embodiment for the preparation of polypeptides according to the invention is the use of peptide mimetics. Mimetics are peptide-containing molecules that mimic elements of protein secondary structure. See, for example, Johnson *et al.*, "Peptide Turn Mimetics" in *BIOTECHNOLOGY AND PHARMACY*, Pezzuto *et al.*, Eds., Chapman and 25 Hall, New York (1993). The underlying rationale behind the use of peptide mimetics is that the peptide backbone of proteins exists chiefly to orient amino acid side chains in such a way as to facilitate molecular interactions, such as those of antibody and antigen. A peptide mimetic is expected to permit molecular interactions similar to the natural molecule.

Successful applications of the peptide mimetic concept have thus far focused on mimetics of  $\beta$ -turns within proteins, which are known to be highly antigenic. Likely  $\beta$ -turn structure within an polypeptide can be predicted by computer-based algorithms as discussed above. Once the component amino acids of the turn are determined, peptide mimetics can 5 be constructed to achieve a similar spatial orientation of the essential elements of the amino acid side chains.

Modification and changes may be made in the structure of a gene and still obtain a functional molecule that encodes a protein or polypeptide with desirable characteristics. 10 The following is a discussion based upon changing the amino acids of a protein to create an equivalent, or even an improved, second-generation molecule. The amino acid changes may be achieved by change the codons of the DNA sequence, according to the following data.

15 For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid substitutions can be made in a 20 protein sequence, and its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the DNA sequences of genes without appreciable loss of their biological utility or activity.

25 In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte & Doolittle, 1982).

**TABLE 1**

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUU	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

It is accepted that the relative hydrophobic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, 5 antibodies, antigens, and the like.

Each amino acid has been assigned a hydrophobic index on the basis of their hydrophobicity and charge characteristics (Kyte & Doolittle, 1982), these are: Isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); 10 methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydrophobic index or score and still result in a protein with similar biological activity, *i.e.*, still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydrophobic indices are within  $\pm 2$  is preferred, those which are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological 25 property of the protein.

As detailed in U.S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0  $\pm$  1); glutamate (+3.0  $\pm$  1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); 30 threonine (-0.4); proline (-0.5  $\pm$  1); alanine (-0.5); histidine \*-0.5); cysteine (-1.0);

methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4).

5 It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent and immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$  is preferred, those that are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

10 As outlined above, amino acid substitutions are generally based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and 15 asparagine; and valine, leucine and isoleucine.

**b. apoB100 Variants**

20 In order to determine the optimal DNA-binding sequences, recombinant fragments of apoB-100 or other apolipoproteins may be used in mobility shift assays or other common protein-DNA interaction assays, including, but not limited to, methylation interference assays, DNase-I footprinting assays, UV-crosslinking assays, Biotin/Streptavidin affinity systems, or screening expression libraries encoding DNA-binding proteins. The recombinant apolipoprotein fragments are expressed by cloning these cDNA fragments in commercially available eukaryotic expression vectors and 25 employing recombinant DNA expression techniques well known to the art.

30 In addition, the recombinant fragments may be mutated by employing site-directed mutagenesis or oligonucleotide-directed mutagenesis techniques in order to improve their affinity for nucleic acids and used either in their original or mutated form. Mutations in the recombinant apolipoprotein fragments may include, but are not limited

to, addition of endosomolytic and/or nuclear localization peptide sequences employing common recombinant DNA technology. The recombinant protein fragments are prebound to the nucleic acids of interest prior to their reassembly into freshly isolated lipoproteins and subsequent transfection. Alternatively, they are reassembled into 5 lipoproteins prior to *in vitro* nucleic acid binding and subsequent transfection. Separation of protein-bound DNA from free DNA may be required prior to transfection and is accomplished by adsorption to nitrocellulose membranes or other common techniques including, but not limited to size-exclusion or density ultracentrifugation.

10        Site specific mutations can be made within the proposed DNA binding motifs or nuclear localization signal sequences of the apolipoproteins described in this invention, in order to improve their homology with known DNA binding motifs (e.g., SREBP-like DNA-binding motifs, ISGF3 $\gamma$ -like DNA-binding motifs) and nuclear localization signal sequences (e.g., NLS from human p53, Ap 1, IGFBP-3, ir, and apo J). Specific mutations 15 in the DNA sequences of steroid regulatory elements (SRE) and IFN-stimulated response elements which affect the DNA-binding affinity of SREBP and ISGF3 $\gamma$ , respectively, have been described (Smith *et al.*, 1990; Briggs *et al.*, 1993; Wang *et al.*, 1993; Veals *et al.*, 1992).

20        Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent proteins or peptides, through specific mutagenesis of the underlying DNA. The technique further provides a ready ability to prepare and test sequence variants, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence change(s) into the DNA. 25 Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to 25 nucleotides in

length is preferred, with about 5 to 10 residues on both sides of the junction of the sequence being altered.

In general, the technique of site-specific mutagenesis is well known in the art. As  
5 will be appreciated, the technique typically employs a bacteriophage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage vectors are commercially available and their use is generally well known to those skilled in the art. Double stranded plasmids are also routinely employed in site directed mutagenesis,  
10 which eliminates the step of transferring the gene of interest from a phage to a plasmid.

In general, site-directed mutagenesis is performed by first obtaining a single-stranded vector, or melting of two strands of a double stranded vector which includes within its sequence a DNA sequence encoding the desired protein. An oligonucleotide primer bearing the desired mutated sequence is synthetically prepared. This primer is then annealed with the single-stranded DNA preparation, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand  
15 bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected that include recombinant vectors bearing the mutated sequence arrangement.  
20

The preparation of sequence variants of the selected gene using site-directed mutagenesis is provided as a means of producing potentially useful species and is not meant to be limiting, as there are other ways in which sequence variants of genes may be obtained. For example, recombinant vectors encoding the desired gene may be treated  
25 with mutagenic agents, such as hydroxylamine, to obtain sequence variants.

## 6. PURIFICATION OF LIPOPROTEINS

The purification of plasma LDL involves obtaining a composition of Lp(a) and subjecting the composition to reductive cleavage in a manner that allows the formation of cleavage products apo (a) and apoB100. These products are then separated to yield purified apo B100. Plasma lipoproteins may be isolated using standard sequential flotation ultracentrifugation methods as described (Schumaker and Puppione, 1986).

### a. Purification of Lp(a)

Lp(a) is known to be made in the liver of primates. The LDL and VLDL in the plasma represents the primary source for the purification of Lp(a). Plasma may be collected from any primate source for the purposes of the invention, or indeed any other source suspected of possessing Lp(a). The Lp(a) component of the plasma can then be separated from other components of the plasma using ultracentrifugational flotation at a density of 1.21 g/mL for 20 hours at 50, 000rpm followed by affinity chromatography using lysine-Sepharose<sup>TM</sup>. Of course, the ultra centrifugational procedure is only exemplary and those of skill in the art will be able to vary them according to the particular equipment and study need without undue experimentation. The plasma may be supplemented with various inhibitors to prevent the Lp(a) from interacting with LDL components of the plasma.

20

Having separated Lp(a) from the other plasma components the Lp(a) sample is purified using affinity chromatography lysine-Sepharose<sup>TM</sup> chromatography. This separation is described in detail in PCT publication WO 97/17371, specifically incorporated herein by reference.

25

In some cases, it is desirable to use a method other than lysine-Sepharose<sup>TM</sup> chromatography for the purification of Lp(a), in such instances other chromatographic methods such FPLC may be employed. Such methods are disclosed in Scanu *et al.*, 1993, incorporated herein by reference, and may be used in conjunction with the present 30 invention to purify apo B100 from Lp(a).

The product purity can be assessed by for example, mobility on, 1% agarose gels, Western blots of SDS PAGE, utilizing anti-LDL antibodies.

5           **b. Isolation of Apo B100 from Lp (a)**

(i)       *using centrifugation*

Following the purification of Lp(a), the apoB100 may be separated from the apo A fraction of the Lpa complex using reductive cleavage.. The purified intact Lp(a) protein is subjected to reductive cleavage wherein a known volume of Lp(a) is incubated with a 10 reductant. Exemplary reductants include homocysteine, N-acetyl cysteine, 2-mercaptopropanoic acid, 3-mercaptopropionate, 2-aminoethanol, dithiothreitol, and DTE.

15           The reaction is incubated at room temperature for 10-20 minutes. This is followed by the addition of an inhibitor to prevent non-covalent, lysine mediated interactions between apo (a) and apoB100.  $\epsilon$ -Aminocaproic acid (EACA) may be used as such an inhibitor. substituted by other lysine analogues, for example, compounds such as trans 4(amino-methyl)-cyclohexanecarboxylic acid, N-acetyl-L-lysine, p-benzylamine sulfonic acid, hexylamine, benzamidine, benzylamine, L-proline. Of course these are only exemplary lysine analogues and those of skill in the art may use other lysine analogues to 20 prevent interaction between apo (a) and apoB100 proteins. The reaction conditions are described in greater detail in PCT publication number WO 97/17371. Of course, the conditions for the separation of apo (a) from the reaction mixture using sucrose density ultracentrifugation is only exemplary, and other methods commonly used by those of skill in the art may be used.

25

(ii)       **Isolation Using Chromatographic Methods**

As an alternative to the above methods for the isolation of apo B100 from Lp(a) chromatographic methods may be utilized as exemplified below.

*Heparin Sepharose™ Chromatography*

Lp(a) may be treated with a reducing agent in the presence of a lysine analogue. For the purposes of this invention the lysine analog is supplied to prevent the interaction of apo (a) with apoB100. The reducing agent is supplied to break the disulfide bond of Lp (a).  
5 Lysine analogs for this invention include but are not limited to compounds such as EACA, trans 4(amino-methyl)-cyclohexanecarboxylic acid, N-acetyl-L-lysine, p-benzylamine sulfonic acid, hexylamine, benzamidine, benzylamine, L-proline or any other lysine analogue known to the artisan skilled in the art may be used. Example of reducing agents that may be used in this invention include, but are not limited to, homocysteine, N-acetyl 10 cysteine, 2-mercaptopropanoate, 2-aminoethanol, dithiothreitol, and DTE.

For example, the mixture of Lp (a), a reducing agent and a lysine analog is incubated for a suitable period of time in a suitable buffer of pH 7.4. A heparin- 15 Sepharose™ column is equilibrated with a suitable buffer containing the lysine analog and the reducing agent. The mixture is applied to the equilibrated column, the column is washed with the same buffer and the first eluate is collected.

The first eluate from the column contains the apo (a) dissociated from Lp (a). The 20 "free" apo (a) is dialyzed against an appropriate buffer. the dialysis product is pure apo (a) that may be freeze dried and stored at -20°C or used immediately. The column is further washed with the buffer for a total of three column volumes followed by 3 volumes of 2M NaCl in the buffer. The high salt concentration serves to dissociate the remaining unreacted Lp(a) and LDL containing apoB100 free of apo (a).

25

*Lysine-Sepharose™ Chromatography*

An alternative to heparin-Sepharose™ chromatography is lysine chromatography. In this type of separation, Lp(a) is treated with a suitable reducing agent and then applied to a lysine Sepharose™ column that has been equilibrated with a suitable buffer of pH 7.4 30 containing the reducing agent. The column is washed with the same buffer and the first

volume of elute is collected. This fraction contains LDL dissociated from apo (a). Further details of this type of chromatography for separating apolipoproteins may be found in PCT Publication WO 97/17371.

5      **7. SCREENING NUCLEIC ACIDS THAT BIND LDL**

Specifically contemplated by the present inventors are chip-based DNA technologies such as those described by Hacia *et al.* (1996) and Shoemaker *et al.* (1996). Chip technologies may be used to present DNA arrays for screening.

10     In a first embodiment, chip technologies may be employed to synthesize a variety of DNAs in order to test for their binding to an LDL with a specific apoB100 binding region. Briefly, these techniques involve quantitative methods for analyzing large numbers of nucleic acids rapidly and accurately. By tagging genes with oligonucleotides or using fixed probe arrays, one can employ chip technology to segregate target molecules as high density arrays and screen these molecules on the basis of hybridization.

15     See also Pease *et al.* (1994); Fodor *et al.* (1991).

20     Thus, the invention may be applied for the screening of nucleic acids that bind to apoB100 containing lipoproteins. The LDL polypeptide or fragment may be either free in solution, fixed to a support, expressed in or on the surface of a cell, for example a bacterial cell. Either the LDL polypeptide or the nucleic acid may be labeled, thereby permitting determining of binding to the DNA molecules.

25     In another embodiment, the assay may measure the inhibition of binding of LDL to a natural or artificial substrate or binding partner. Competitive binding assays can be performed in which one of the agents (LDL, binding partner or compound) is labeled. Usually, the polypeptide will be the labeled species. One may measure the amount of free label versus bound label to determine binding or inhibition of binding.

Another technique for high throughput screening of compounds is described in WO 84/03564. Large numbers of small test nucleic acids (test compounds) are synthesized on a solid substrate, such as plastic pins or some other surface. Similarly, test compounds of the present invention are reacted with LDL and washed. Bound 5 polypeptide is detected by various methods.

In an alternative embodiment, the invention may be applied for the screening for variants of apoB100 containing lipoproteins to determine a greater or lesser affinity for a particular type of nucleic acid. These screening methods would be similar to those 10 described above, except that the LDL peptide variants will be presented as an array with the nucleic acid binding regions being used to probe the array. Currently, one of the most widely used approaches for screening polypeptide libraries is to display polypeptides on the surface of filamentous bacteriophage (Smith, 1991; Smith, 1992). Ladner *et al.*, (U.S Patent No 5,403,484, specifically incorporated herein by reference) reported the display 15 of proteins on the outer surface of a chosen bacterial cell, spore or phage, in order to identify and characterize binding proteins.

In an alternative embodiment, purified apoB100 or DNA-binding fragments thereof can be coated directly onto plates for use in the screening techniques. 20 Alternatively, antibodies to the polypeptide can be used to immobilize the polypeptide to a solid phase. Also, fusion proteins containing a DNA binding region (preferably a terminal region) may be used to link peptides to a solid phase. Once linked, randomly sheared genomic DNA, transcripts or randomly generated oligomers may be contacted with the bound peptides. Any bound nucleic acid fragments can be identified by PCR 25 using random primers if they are large enough. In the case where random oligomers are used, the oligomers, in addition to the random region, may comprise built in primer binding sites that can be used to amplify an intervening random region, thereby identifying the region binding to apoB100.

Thus, using the technologies described herein, it will be possible for one of skill in the art to screen for and isolate a variety of nucleic acids that bind to apoB100 and variants of apoB100 that exhibit nucleic acid binding capacity, including increased or decreased binding as compared to wild-type apoB100.

5

## 8. LDL-DNA COMPLEX FORMATION

In particular aspects of the present invention, lipoproteins are employed in order to transport DNA into cell *in vitro* and *in vivo*. In the present invention, optimal DNA/LDL binding has been established. In particular embodiments a 1:1 ratio of 10 DNA:LDL protein molar ratio of 1:1 are incubated at 37 °C for 30 min in a buffered solution. An exemplary buffer may be 50 mM Tris-HCl at pH 7.4 containing 150 mM NaCl, and 10 mM MgCl<sub>2</sub>. The concentrations of DNA and LDL protein may range from the pmolar range to the μmolar range. In a preferred embodiment, 0.39 pmole DNA are incubated with 0.39 pmole LDL-protein.

15

The incubation conditions may be altered to increase or decrease the efficiency of DNA/LDL binding. For example the incubation may occur at temperatures ranging from 4°C to 50°C, thus it is contemplated that the reaction mixture may be incubated at 4°C, 6°C, 8°C 10°C, 12°C, 14°C, 16°C, 18°C, 20°C, 22°C, 24°C, 26°C, 28°C, 30°C, 32°C, 20 34°C, 36°C, 38°C, 40°C, 42°C, 44°C, 46°C, 48°C, 50°C.

The time of incubation may be varied from as little as 10 minutes to as long as 5 hours. Thus it is well within the skill of one in the art to incubate the mixture for varying degrees of time.

25

Other embodiments contemplate varying the concentration of MgCl<sub>2</sub> in the media. Thus the MgCl<sub>2</sub> concentration may vary from 1mM to 100 mM. Thus, it is contemplated that the reaction mixture contains 5mM MgCl<sub>2</sub>, 10mM MgCl<sub>2</sub>, 12mM MgCl<sub>2</sub>, 15mM MgCl<sub>2</sub>, 20mM MgCl<sub>2</sub>, 30mM MgCl<sub>2</sub>, 35mM MgCl<sub>2</sub>, 40mM MgCl<sub>2</sub>, 50mM

MgCl<sub>2</sub>, 60mM MgCl<sub>2</sub>, 65mM MgCl<sub>2</sub>, 70mM MgCl<sub>2</sub>, 80mM MgCl<sub>2</sub>, 90mM MgCl<sub>2</sub>, or 100mM MgCl<sub>2</sub>.

## 9. GENE DELIVERY AND EXPRESSION IN EUKARYOTIC CELLS

5 The gene delivery system of the instant invention can be used to express any gene of interest in eukaryotic cells. The gene or its cDNA sequence is cloned into a plasmid containing the specific lipoprotein binding sequences (including, but not limited to SRE, E/C, FAS) and/or any eukaryotic regulatory sequence (for example, but not limited to HCMV, or tyrosine kinase promoter region) using DNA cloning techniques well known to the art. The orientation, number and location of the lipoprotein binding sequences may 10 vary within the nucleic acid vector, but should not interrupt the protein coding sequence of the gene of interest.

15 The gene delivery system of the instant invention (see FIG. 15) can be used to transfect eukaryotic cells either *in vivo* or *in vitro* with any expression vector containing one or more of the aforementioned lipoprotein binding sequences. Expression vectors are designed using recombinant DNA cloning techniques known to the art and generally include five components linked in the following 5' to 3' orientation: i) an eukaryotic promoter sequence, 2) a sequence encoding a 5' untranslated RNA (UTR) which may 20 include a first intron sequence followed by a consensus Kozak sequence and an initiation ATG, 3) a protein coding sequence, 4) a 3' UTR, and 5) a cognate transcription terminator sequence.

25 Lipoproteins are isolated from blood in a manner similar to the previously described procedures (see, Example 1) and bound to the nucleic acids of interest in a manner similar to the previously described DNA binding protocol (see, Example 2). Separation of protein-bound DNA from free DNA may be required prior to transfection and can be accomplished by adsorption to nitrocellulose membranes or other techniques well known to the art including, but not limited to size-exclusion or density 30 ultracentrifugation.

### a) Control Regions

In order for the gene delivery system of the present invention to effect expression of a transcript encoding a selected gene, the polynucleotides encoding these genes will be under the transcriptional control of a promoter. A “promoter” refers to a DNA sequence recognized by the synthetic machinery of the host cell, or introduced synthetic machinery, that is required to initiate the specific transcription of a gene. The phrase “under transcriptional control” means that the promoter is in the correct location in relation to the polynucleotide to control RNA polymerase initiation and expression of the polynucleotide.

The term promoter will be used here to refer to a group of transcriptional control modules that are clustered around the initiation site for RNA polymerase II. Much of the thinking about how promoters are organized derives from analyses of several viral promoters, including those for the HSV thymidine kinase (tk) and SV40 early transcription units. These studies, augmented by more recent work, have shown that promoters are composed of discrete functional modules, each consisting of approximately 7-20 bp of DNA, and containing one or more recognition sites for transcriptional activator or repressor proteins.

20

At least one module in each promoter functions to position the start site for RNA synthesis. The best known example of this is the TATA box, but in some promoters lacking a TATA box, such as the promoter for the mammalian terminal deoxynucleotidyl transferase gene and the promoter for the SV40 late genes, a discrete element overlying the start site itself helps to fix the place of initiation.

Additional promoter elements regulate the frequency of transcriptional initiation. Typically, these are located in the region 30-110 bp upstream of the start site, although a number of promoters have recently been shown to contain functional elements downstream of the start site as well. The spacing between promoter elements frequently

is flexible, so that promoter function is preserved when elements are inverted or moved relative to one another. In the tk promoter, the spacing between promoter elements can be increased to 50 bp apart before activity begins to decline. Depending on the promoter, it appears that individual elements can function either cooperatively or independently to 5 activate transcription.

The particular promoter that is employed to control the expression of a therapeutic gene is not believed to be critical, so long as it is capable of expressing the polynucleotide in the targeted cell. Thus, where a human cell is targeted, it is preferable to position the 10 polynucleotide coding region adjacent to and under the control of a promoter that is capable of being expressed in a human cell. Generally speaking, such a promoter might include either a human or viral promoter.

In preferred embodiments, the human cytomegalovirus (CMV) immediate early 15 gene promoter, the SV40 early promoter and the Rous sarcoma virus long terminal repeat can be used to obtain high-level expression of the polynucleotide of interest. The use of other viral or mammalian cellular or bacterial phage promoters which are well-known in the art to achieve expression of polynucleotides is contemplated as well, provided that the levels of expression are sufficient to produce a growth inhibitory effect.

20

By employing a promoter with well-known properties, the level and pattern of expression of a polynucleotide following transfection can be optimized. For example, selection of a promoter which is active in specific cells, such as tyrosinase (melanoma), alpha-fetoprotein and albumin (liver tumors), CC10 (lung tumor) and prostate-specific 25 antigen (prostate tumor) will permit tissue-specific expression of the therapeutic gene.

Enhancers were originally detected as genetic elements that increased transcription from a promoter located at a distant position on the same molecule of DNA. This ability to act over a large distance had little precedent in classic studies of 30 prokaryotic transcriptional regulation. Subsequent work showed that regions of DNA

with enhancer activity are organized much like promoters. That is, they are composed of many individual elements, each of which binds to one or more transcriptional proteins.

5        The basic distinction between enhancers and promoters is operational. An enhancer region as a whole must be able to stimulate transcription at a distance; this need not be true of a promoter region or its component elements. On the other hand, a promoter must have one or more elements that direct initiation of RNA synthesis at a particular site and in a particular orientation, whereas enhancers lack these specificities. Promoters and enhancers are often overlapping and contiguous, often seeming to have a  
10      very similar modular organization.

Additionally, any promoter/enhancer combination (as per the Eukaryotic Promoter Data Base EPDB) could be used to drive expression of a particular construct. Use of a T3, T7 or SP6 cytoplasmic expression system is another possible embodiment.  
15      Eukaryotic cells can support cytoplasmic transcription from certain bacteriophage promoters if the appropriate bacteriophage polymerase is provided, either as part of the delivery complex or as an additional genetic expression vector.

According to the present invention, a number of different promoters are required.  
20      It is contemplated that these promoters may be the same or different, but the selection of particular promoters for particular uses may be advantageous.

**b)      IRES**

In certain embodiments of the invention, the use of internal ribosome binding site  
25      (IRES) elements may prove advantageous in accordance with the present invention. These elements are used to create multigene, or polycistronic, messages. IRES elements are able to bypass the ribosome scanning model of 5' methylated Cap dependent translation and begin translation at internal sites (Pelletier and Sonenberg, 1988). IRES elements from two members of the picornavirus family (polio and encephalomyocarditis)  
30      have been described (Pelletier and Sonenberg, 1988), as well an IRES from a mammalian

message (Macejak and Sarnow, 1991). IRES elements can be linked to heterologous open reading frames. Multiple open reading frames can be transcribed together, each separated by an IRES, creating polycistronic messages. By virtue of the IRES element, each open reading frame is accessible to ribosomes for efficient translation. Multiple 5 genes can be efficiently expressed using a single promoter/enhancer to transcribe a single message.

Any heterologous open reading frame can be linked to IRES elements. This includes genes for secreted proteins, multi-subunit proteins, encoded by independent 10 genes, intracellular or membrane-bound proteins and selectable markers. In this way, expression of several proteins can be simultaneously engineered into a cell with a single construct and a single selectable marker.

In addition, it may be desirable to include polyadenylation signals in the vectors. 15 These signals serve to terminate transcription and to stabilize mRNA transcripts produced from the vectors. A preferred polyadenylation signal is an SV40 polyadenylation signal.

**c) Genes**

The present invention contemplates the use of a variety of different genes inserted 20 into the SV40 vector. For example, genes encoding enzymes, hormones, cytokines, oncogenes, receptors, tumor suppressors, transcription factors, drug selectable markers, toxins and various antigens are contemplated as suitable genes for use according to the present invention. In addition, antisense constructs derived from oncogenes are other “genes” of interest according to the present invention.

25

A common gene currently being used in many gene therapy trials is p53, which currently is recognized as a tumor suppressor gene. High levels of mutant p53 have been found in many cells transformed by chemical carcinogenesis, ultraviolet radiation, and several viruses. The p53 gene is a frequent target of mutational inactivation in a wide 30 variety of human tumors and is already documented to be the most frequently-mutated

gene in common human cancers. It is mutated in over 50% of human NSCLC (Hollstein *et al.*, 1991) and in a wide spectrum of other tumors. Overexpression of wild-type p53 has been shown in some cases to be anti-proliferative in human tumor cell lines. Thus, p53 can act as a negative regulator of cell growth (Weinberg, 1991) and may directly 5 suppress uncontrolled cell growth or indirectly activate genes that suppress this growth. It has also been reported that transfection of DNA encoding wild-type p53 into cancer cell lines restores growth suppression control in such cells (Casey *et al.*, 1991; Takahasi *et al.*, 1992). It is thus proposed that the treatment of p53-associated cancers with wild 10 type p53 in the compositions of the present invention will reduce the number of malignant cells or their growth rate.

15 p16<sup>INK4</sup> belongs to a newly described class of CDK-inhibitory proteins that also includes p16<sup>B</sup>, p21<sup>WAF1</sup>, and p27<sup>KIP1</sup>. The p16<sup>INK4</sup> gene maps to 9p21, a chromosome region frequently deleted in many tumor types. Homozygous deletions and mutations of the p16<sup>INK4</sup> gene are frequent in human tumor cell lines. This evidence suggests that the p16<sup>INK4</sup> gene is a tumor suppressor gene. This interpretation has been challenged, however, by the observation that the frequency of the p16<sup>INK4</sup> gene alterations is much lower in primary uncultured tumors than in cultured cell lines (Caldas *et al.*, 1994; Cheng *et al.*, 1994; Hussussian *et al.*, 1994; Kamb *et al.*, 1994; Kamb *et al.*, 1994; Mori *et al.*, 20 1994; Okamoto *et al.*, 1994; Nobori *et al.*, 1995; Orlow *et al.*, 1994; Arap *et al.*, 1995). Restoration of wild-type p16<sup>INK4</sup> function by transfection with a plasmid expression vector reduced colony formation by some human cancer cell lines (Okamoto, 1994; Arap, 1995).

25 Cell adhesion molecules, or CAM's are known to be involved in a complex network of molecular interactions that regulate organ development and cell differentiation (Edelman, 1985). Recent data indicate that aberrant expression of CAM's maybe involved in the tumorigenesis of several neoplasms; for example, decreased expression of E-cadherin, which is predominantly expressed in epithelial cells, is associated with the 30 progression of several kinds of neoplasms (Edelman and Crossin, 1991; Frixen *et al.*,

1991; Bussemakers *et al.*, 1992; Matsura *et al.*, 1992; Umbas *et al.*, 1992). Also, 5 Giancotti and Ruoslahti (1990) demonstrated that increasing expression of  $\alpha_5\beta_1$  integrin by gene transfer can reduce tumorigenicity of Chinese hamster ovary cells *in vivo*. C-CAM now has been shown to suppress tumors growth *in vitro* and *in vivo*. Thus, the compositions of the present invention can be employed to mediated C-CAM suppression of tumor cell growth.

Other tumor suppressors that may be employed according to the present invention include RB, APC, DCC, NF-1, NF-2, WT-1, MEN-I, MEN-II, zac1, p73, VHL, 10 MMAC1, FCC and MCC. Inducers of apoptosis, such as Bax, Bak, Bcl-X<sub>s</sub>, Bik, Bid, Harakiri, Ad E1B, Bad and ICE-CED3 proteases, similarly could find use according to the present invention.

Various enzyme genes are of interest according to the present invention. Such 15 enzymes include cytosine deaminase, hypoxanthine-guanine phosphoribosyltransferase, galactose-1-phosphate uridyltransferase, phenylalanine hydroxylase, glucocerbrosidase, sphingomyelinase,  $\alpha$ -L-iduronidase, glucose-6-phosphate dehydrogenase, HSV thymidine kinase and human thymidine kinase.

20 In another example, the expression vector may include a nucleotide sequence encoding for functional apolipoprotein A-I for the prevention or treatment of atherosclerosis. Atherosclerosis is a disease that is characterized by the development of atherosclerotic lesions which contain cholesterol esters and other lipids that are derived from the blood circulation. The plasma concentration of HDL is inversely correlated with 25 the risk for development of atherosclerosis. HDL present in the blood circulation take up free cholesterol from extrahepatic cells which through the action of LCAT (lecithin-cholesterol acyltransferase) is converted to cholesterol esters and stored in the core of the HDL particles. The HDL cholesterol esters are transported either directly or indirectly via transfer to triglyceride rich lipoproteins (*i.e.*, VLDL, IDL, LDL) to the liver by a 30 process called "reverse cholesterol transport". Reverse cholesterol transport is of great

importance for maintaining cholesterol homeostasis since the liver is the major organ for cholesterol excretion from the body *via* bile acids. Apo A-I is the major protein constituent of HDL and a cofactor LCAT. Therefore, increasing the plasma concentration of apo A-I containing HDL can increase the reverse cholesterol transport and reduce the risk for atherosclerosis.

Hormones are another group of gene that may be used in the SV40 vectors described herein. Included are growth hormone, prolactin, placental lactogen, luteinizing hormone, follicle-stimulating hormone, chorionic gonadotropin, thyroid-stimulating hormone, leptin, adrenocorticotropin (ACTH), angiotensin I and II,  $\beta$ -endorphin,  $\beta$ -melanocyte stimulating hormone ( $\beta$ -MSH), cholecystokinin, endothelin I, galanin, gastric inhibitory peptide (GIP), glucagon, insulin, lipotropins, neuropeptides, somatostatin, calcitonin, calcitonin gene related peptide (CGRP),  $\beta$ -calcitonin gene related peptide, hypercalcemia of malignancy factor (1-40), parathyroid hormone-related protein (107-139) (PTH-rP), parathyroid hormone-related protein (107-111) (PTH-rP), glucagon-like peptide (GLP-1), pancreastatin, pancreatic peptide, peptide YY, PHM, secretin, vasoactive intestinal peptide (VIP), oxytocin, vasopressin (AVP), vasotocin, enkephalinamide, metorphinamide, alpha melanocyte stimulating hormone (alpha-MSH), atrial natriuretic factor (5-28) (ANF), amylin, amyloid P component (SAP-1), corticotropin releasing hormone (CRH), growth hormone releasing factor (GHRH), luteinizing hormone-releasing hormone (LHRH), neuropeptide Y, substance K (neurokinin A), substance P and thyrotropin releasing hormone (TRH).

Other classes of genes that are contemplated to be inserted into the SV40 vectors of the present invention include interleukins and cytokines. Interleukin 1 (IL-1), IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11 IL-12, GM-CSF and G-CSF.

Other therapeutics genes might include genes encoding antigens such as viral antigens, bacterial antigens, fungal antigens or parasitic antigens. Viruses include picornavirus, coronavirus, togavirus, flavivirus, rhabdovirus, paramyxovirus,

orthomyxovirus, bunyavirus, arenavirus, reovirus, retrovirus, papovavirus, parvovirus, herpesvirus, poxvirus, hepadnavirus, and spongiform virus. Preferred viral targets include influenza, herpes simplex virus 1 and 2, measles, small pox, polio or HIV. Pathogens include trypanosomes, tapeworms, roundworms, helminths, . Also, tumor 5 markers, such as fetal antigen or prostate specific antigen, may be targeted in this manner. Preferred examples include HIV env proteins and hepatitis B surface antigen. Administration of a vector according to the present invention for vaccination purposes would require that the vector-associated antigens be sufficiently non-immunogenic to enable long term expression of the transgene, for which a strong immune response would 10 be desired. Preferably, vaccination of an individual would only be required infrequently, such as yearly or biennially, and provide long term immunologic protection against the infectious agent.

15 In yet another embodiment, the heterologous gene may include a single-chain antibody. Methods for the production of single-chain antibodies are well known to those of skill in the art. The skilled artisan is referred to U.S. Patent No. 5,359,046, (incorporated herein by reference) for such methods. A single chain antibody is created by fusing together the variable domains of the heavy and light chains using a short peptide linker, thereby reconstituting an antigen binding site on a single molecule.

20

25 Single-chain antibody variable fragments (Fvs) in which the C-terminus of one variable domain is tethered to the N-terminus of the other via a 15 to 25 amino acid peptide or linker, have been developed without significantly disrupting antigen binding or specificity of the binding (Bedzyk *et al.*, 1990; Chaudhary *et al.*, 1990). These Fvs lack the constant regions (Fc) present in the heavy and light chains of the native antibody.

Antibodies to a wide variety of molecules are contemplated, such as oncogenes, toxins, hormones, enzymes, viral or bacterial antigens, transcription factors or receptors.

30

**d. Antisense**

The instant invention can be used to transfect eukaryotic cells with ribonucleotide sequences including anti-sense RNA and ribozymes, that function to inhibit the translation of any mRNA of interest, either by direct binding (to the mRNA of interest), or blocking deoxyribonucleic acid (DNA) coding sequences preventing transcription.

5

Anti-sense RNA inhibits the translation of mRNA by direct binding to the mRNA of interest and preventing protein translation, either by inhibition of ribosome binding or the translocation of the targeted mRNA molecule which then becomes more susceptible to nuclease degradation.

10

Antisense methodology takes advantage of the fact that nucleic acids tend to pair with "complementary" sequences. By complementary, it is meant that polynucleotides are those which are capable of base-pairing according to the standard Watson-Crick complementarity rules. That is, the larger purines will base pair with the smaller pyrimidines to form combinations of guanine paired with cytosine (G:C) and adenine paired with either thymine (A:T) in the case of DNA, or adenine paired with uracil (A:U) in the case of RNA. Inclusion of less common bases such as inosine, 5-methylcytosine, 6-methyladenine, hypoxanthine and others in hybridizing sequences does not interfere with pairing. Oncogenes such as *ras*, *myc*, *neu*, *raf*, *erb*, *src*, *fms*, *jun*, *trk*, *ret*, *gsp*, *hst*, *bcl* and *abl* also are suitable targets for antisense constructs.

Targeting double-stranded (ds) DNA with polynucleotides leads to triple-helix formation; targeting RNA will lead to double-helix formation. Antisense polynucleotides, when introduced into a target cell, specifically bind to their target polynucleotide and interfere with transcription, RNA processing, transport, translation and/or stability. Antisense RNA constructs, or DNA encoding such antisense RNA's, may be employed to inhibit gene transcription or translation or both within a host cell, either in vitro or in vivo, such as within a host animal, including a human subject.

Antisense constructs may be designed to bind to the promoter and other control regions, exons, introns or even exon-intron boundaries of a gene. It is contemplated that the most effective antisense constructs will include regions complementary to intron/exon splice junctions. Thus, it is proposed that a preferred embodiment includes an antisense construct with complementarity to regions within 50-200 bases of an intron-exon splice junction. It has been observed that some exon sequences can be included in the construct without seriously affecting the target selectivity thereof. The amount of exonic material included will vary depending on the particular exon and intron sequences used. One can readily test whether too much exon DNA is included simply by testing the constructs *in vitro* to determine whether normal cellular function is affected or whether the expression of related genes having complementary sequences is affected.

As stated above, "complementary" or "antisense" means polynucleotide sequences that are substantially complementary over their entire length and have very few base mismatches. For example, sequences of fifteen bases in length may be termed complementary when they have complementary nucleotides at thirteen or fourteen positions. Naturally, sequences which are completely complementary will be sequences which are entirely complementary throughout their entire length and have no base mismatches. Other sequences with lower degrees of homology also are contemplated. For example, an antisense construct which has limited regions of high homology, but also contains a non-homologous region (e.g., ribozyme) could be designed. These molecules, though having less than 50% homology, would bind to target sequences under appropriate conditions.

It may be advantageous to combine portions of genomic DNA with cDNA or synthetic sequences to generate specific constructs. For example, where an intron is desired in the ultimate construct, a genomic clone will need to be used. The cDNA or a synthesized polynucleotide may provide more convenient restriction sites for the remaining portion of the construct and, therefore, would be used for the rest of the sequence.

e. **Ribozymes**

Ribozymes are RNA molecules that catalyze the specific cleavage of RNA. Ribozyme activity is mediated through the hybridization of the ribozyme molecule to a specific sequence in the target RNA, followed by the endonucleolytic cleavage of the target RNA within that sequence. Potential RNA cleavage sites can be identified by searching for specific ribonucleotide sequences that include sequences such as GUU, GUC, and GUA within the target RNA. Hammerhead motif ribozyme molecules can then be designed that contain short RNA sequences (15-25 ribonucleotides) that are complementary to the region including the cleavage site of the target RNA.

Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cook, 1987; Gerlach *et al.*, 1987; Forster and Symons, 1987). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cook *et al.*, 1981; Michel and Westhof, 1990; Reinholt-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cook *et al.*, 1981). For example, U.S. Patent No. 5,354,855 reports that certain ribozymes can act as endonucleases with a sequence specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon *et al.*, 1991; Sarver *et al.*, 1990). Recently, it was reported that ribozymes elicited genetic changes in some cell lines to which they were applied; the altered genes included the oncogenes H-ras, c-fos and genes of HIV. Most of this work

involved the modification of a target mRNA, based on a specific mutant codon that is cleaved by a specific ribozyme.

Since the secondary structure of both target RNA as well as the anti-sense RNA is of great importance for the hybridization of both molecules, the predicted structural features can be analyzed and RNase protection assays can be used to determine hybridization efficiency. Anti-sense RNA and ribozymes can be synthesized employing chemical nucleic acid synthesis techniques well known to the art (*i.e.*, solid phase phosphoromidite synthesis) or the RNA molecules can be produced by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA. DNA sequences encoding ribozymes or anti-sense RNA may be incorporated into an expression vector. The expression vector may be prebound to purified plasma lipoprotein fractions prior to transfection into eukaryotic cells.

**15 f. Self-initiating and self-sustaining gene expression systems**

The invention gene delivery system can also be used to deliver self-initiating and self-sustaining gene expression systems. Self-initiating and self-sustaining gene expression systems may be constructed by binding a RNA polymerase to a DNA construct *in vitro* prior to the introduction of the polynucleotide into the cell as described by Wagner *et al.* (U.S. Patent No. #5,591,601). The RNA polymerase is bound to a DNA construct containing a cognate promoter of the RNA polymerase operably linked to a DNA sequence encoding for the RNA polymerase.

25 The expression of functional RNA polymerase in turn enables the expression of any gene of interest that contains a cognate promoter sequence recognized by the same RNA polymerase in eukaryotic host cells. DNA sequences encoding for both RNA polymerase and gene product of interest (*i.e.*, protein of interest) may be contained within the same gene expression system. The gene expression system may be prebound to purified plasma lipoprotein fractions prior to transfection into eukaryotic cells.

30

g. **Delivery of DNA to cells *in vivo***

The invention gene delivery system can also be used to deliver DNA to cells *in vivo*. An expression vector containing the polynucleotide sequences of the gene of interest (e.g., reporter gene or a healthy copy of a defective gene) is prebound to LDL according to the protocols described herein. This DNA-LDL complex is then introduced into an organism for example, a rat, mouse or human by, for example, intravenous injection. At varying times post-injection, LDL is isolated from the blood and probed for DNA sequences of the type that were prebound to the LDL using standard molecular biological techniques such as, but not limited to, Southern blot hybridization or PCR™.

10

The LDL also can be immunoprecipitated with anti-LDL antibodies and then probed for specific DNA sequences bound to it. In order to determine cellular internalization and/or integration of the reporter gene sequences into the genomic DNA of cells of different tissues, total genomic DNA can be isolated from various tissues (according to standard molecular biology techniques) and probed for the presence of the reporter gene sequences using specific polynucleotide probes in PCR™ or Southern blot hybridization techniques. In addition, total cellular RNA can be isolated from various different tissues using standard molecular biology techniques and probed for the presence of specific mRNA encoded for by the reporter gene polynucleotide sequences using specific antisense polynucleotide probes in Northern blot hybridization techniques or ribonuclease (RNase) protection assays.

20  
25  
Expression of a functional protein encoded for by the gene of interest in different tissues can be analyzed using techniques well known to the art, such as, Western blot hybridization of cellular protein extracts with antibodies that bind specifically to the reporter gene product (i.e., protein of interest) or direct detection of intracellular fluorescence (e.g., when reporter genes are used that encode for blue or green fluorescent proteins (e.g., GFP from Clontech Inc.).

Several non-viral methods for the transfer of a DNA-LDL complex of the present invention into cultured mammalian cells also are contemplated by the present invention. These include calcium phosphate precipitation (Graham and Van Der Eb, 1973; Chen and Okayama, 1987; Rippe *et al.*, 1990) DEAE-dextran (Gopal, 1985), electroporation (Tur-Kaspa *et al.*, 1986; Potter *et al.*, 1984), direct microinjection (Harland and Weintraub, 1985), DNA-loaded liposomes (Nicolau and Sene, 1982; Fraley *et al.*, 1979) and lipofectamine-DNA complexes, cell sonication (Fechheimer *et al.*, 1987), gene bombardment using high velocity microprojectiles (Yang *et al.*, 1990), and receptor-mediated transfection (Wu and Wu, 1987; Wu and Wu, 1988). Some of these techniques 5 may be successfully adapted for *in vivo* or *ex vivo* use.  
10

Once the DNA-LDL complex has been delivered into the cell, the nucleic acid encoding the gene of interest may be positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the gene may be stably integrated into the 15 genome of the cell. This integration may be in the cognate location and orientation via homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or "episomes" encode sequences sufficient to permit maintenance and 20 replication independent of or in synchronization with the host cell cycle. How the DNA-LDL complex is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of DNA molecule bound to the LDL.

In one embodiment of the invention, the DNA-LDL complex may simply consist 25 of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell membrane. This is particularly applicable for transfer *in vitro* but it may be applied to *in vivo* use as well. Dubensky *et al.*, (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn 30 mice demonstrating active viral replication and acute infection. Benvenisty and Neshif

(1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes. It is envisioned that DNA encoding a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

5

Another embodiment of the invention for transferring a naked DNA-LDL complex into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several 10 devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

15

Selected organs including the liver, skin, and muscle tissue of rats and mice have been bombarded *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e.*, *ex vivo* treatment. Again, DNA encoding a particular gene may be delivered via this method and still be incorporated by the present invention.

20

In a further embodiment of the invention, the DNA-LDL complex may be entrapped in a liposome. Liposomes are vesicular structures characterized by a phospholipid bilayer membrane and an inner aqueous medium. Multilamellar liposomes have multiple lipid layers separated by aqueous medium. They form spontaneously when phospholipids are suspended in an excess of aqueous solution. The lipid components undergo self-rearrangement before the formation of closed structures and entrap water and dissolved solutes between the lipid bilayers (Ghosh and Bachhawat, 1991). Also contemplated are lipofectamine-DNA complexes.

5           Liposome-mediated nucleic acid delivery and expression of foreign DNA *in vitro* has been very successful. Wong *et al.*, (1980) demonstrated the feasibility of liposome-mediated delivery and expression of foreign DNA in cultured chick embryo, HeLa and hepatoma cells. Nicolau *et al.*, (1987) accomplished successful liposome-mediated gene transfer in rats after intravenous injection.

In certain embodiments of the invention, the liposome may be complexed with a hemagglutinating virus (HVJ). This has been shown to facilitate fusion with the cell membrane and promote cell entry of liposome-encapsulated DNA (Kaneda *et al.*, 1989).  
10           In other embodiments, the liposome may be complexed or employed in conjunction with nuclear non-histone chromosomal proteins (HMG-1) (Kato *et al.*, 1991). In yet further embodiments, the liposome may be complexed or employed in conjunction with both HVJ and HMG-1. In that such expression constructs have been successfully employed in transfer and expression of nucleic acid *in vitro* and *in vivo*, then they are applicable for  
15           the present invention. Where a bacterial promoter is employed in the DNA construct, it also will be desirable to include within the liposome an appropriate bacterial polymerase.

20           Other DNA-LDL complexes which can be employed to deliver a nucleic acid encoding a particular gene into cells are receptor-mediated delivery vehicles. These take advantage of the selective uptake of macromolecules by receptor-mediated endocytosis in almost all eukaryotic cells. Because of the cell type-specific distribution of various receptors, the delivery can be highly specific (Wu and Wu, 1993).

25           Receptor-mediated gene targeting vehicles generally consist of two components: a cell receptor-specific ligand and a DNA-binding agent. Several ligands have been used for receptor-mediated gene transfer. The most extensively characterized ligands are asialoorosomucoid (ASOR) (Wu and Wu, 1987) and transferrin (Wagner *et al.*, 1990). Recently, a synthetic neoglycoprotein, which recognizes the same receptor as ASOR, has been used as a gene delivery vehicle (Ferkol *et al.*, 1993; Perales *et al.*, 1994) and

epidermal growth factor (EGF) has also been used to deliver genes to squamous carcinoma cells (Myers, EPO 0273085).

In other embodiments, the delivery vehicle may comprise a ligand and a 5 liposome. For example, Nicolau *et al.*, (1987) employed lactosyl-ceramide, a galactose-terminal asialganglioside, incorporated into liposomes and observed an increase in the uptake of the insulin gene by hepatocytes. Thus, it is feasible that a nucleic acid encoding 10 a particular gene also may be specifically delivered into a cell type such as lung, epithelial or tumor cells, by any number of receptor-ligand systems with or without liposomes. For example, epidermal growth factor (EGF) may be used as the receptor for mediated delivery of a nucleic acid encoding a gene in many tumor cells that exhibit upregulation of EGF receptor. Mannose can be used to target the mannose receptor on liver cells. Also, antibodies to CD5 (CLL), CD22 (lymphoma), CD25 (T-cell leukemia) 15 and MAA (melanoma) can similarly be used as targeting moieties.

15

In certain embodiments, gene transfer may more easily be performed under *ex vivo* conditions. *Ex vivo* gene therapy refers to the isolation of cells from an animal, the delivery of a nucleic acid into the cells *in vitro*, and then the return of the modified cells back into an animal. This may involve the surgical removal of tissue/organs from an 20 animal or the primary culture of cells and tissues. Anderson *et al.*, U.S. Patent 5,399,346, and incorporated herein in its entirety, disclose *ex vivo* therapeutic methods.

## 10. PHARMACEUTICAL

The gene delivery system of the instant invention can be administered *in vivo* in 25 various ways including, but not limited to, intravenous, pharyngeal, epidermal, intramuscular, intraperitoneal (IP), nasal, and/or rectal. The gene delivery system of the instant invention can also be used for *in vitro* transfections of eukaryotic cell types which possess specific lipoprotein receptors on their cytoplasmic membranes, but is not limited to these cell types.

30

Pharmaceutical products that may spring from the current invention may comprise naked polynucleotide containing single or multiple copies of the specific nucleotide sequences that bind to specific DNA-binding sites of the apolipoproteins present on plasma lipoproteins as described in the current invention. The polynucleotide may 5 encode a biologically active peptide, antisense RNA, or ribozyme and will be provided in a physiologically acceptable administrable form.

Another pharmaceutical product that may spring from the current invention may comprise a highly purified plasma lipoprotein fraction, isolated according to the 10 methodology, described herein from either the patients blood or other source, and a polynucleotide containing single or multiple copies of the specific nucleotide sequences that bind to specific DNA-binding sites of the apolipoproteins present on plasma lipoproteins, prebound to the purified lipoprotein fraction in a physiologically acceptable, administrable form.

15

Yet another pharmaceutical product may comprise a highly purified plasma lipoprotein fraction which contains recombinant apolipoprotein fragments containing single or multiple copies of specific DNA-binding motifs, prebound to a polynucleotide containing single or multiple copies of the specific nucleotide sequences, in a 20 physiologically acceptable administrable form. Yet another pharmaceutical product may comprise a highly purified plasma lipoprotein fraction which contains recombinant apolipoprotein fragments containing single or multiple copies of specific DNA-binding motifs, prebound to a polynucleotide containing single or multiple copies of the specific nucleotide sequences, in a physiologically acceptable administrable form.

25

The dosage to be administered depends to a great extent on the body weight and physical condition of the subject being treated as well as the route of administration and frequency of treatment. A pharmaceutical composition comprising the naked polynucleotide prebound to a highly purified lipoprotein fraction may be administered in 30 amounts ranging from 1 µg to 1 mg polynucleotide and 1 µg to 100 mg protein.

Administration of the therapeutic virus particle to a patient will follow general protocols for the administration of chemotherapeutics, taking into account the toxicity, if any, of the vector. It is anticipated that the treatment cycles would be repeated as necessary. It also is contemplated that various standard therapies, as well as surgical intervention, may be applied in combination with the described gene therapy.

Where clinical application of a gene therapy is contemplated, it will be necessary to prepare the complex as a pharmaceutical composition appropriate for the intended application. Generally this will entail preparing a pharmaceutical composition that is essentially free of pyrogens, as well as any other impurities that could be harmful to humans or animals. One also will generally desire to employ appropriate salts and buffers to render the complex stable and allow for complex uptake by target cells.

Aqueous compositions of the present invention comprise an effective amount of the compound, dissolved or dispersed in a pharmaceutically acceptable carrier or aqueous medium. Such compositions can also be referred to as inocula. The phrases "pharmaceutically or pharmacologically acceptable" refer to molecular entities and compositions that do not produce an adverse, allergic or other untoward reaction when administered to an animal, or a human, as appropriate. As used herein, "pharmaceutically acceptable carrier" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients also can be incorporated into the compositions.

The compositions of the present invention may include classic pharmaceutical preparations. Dispersions also can be prepared in glycerol, liquid polyethylene glycols,

and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

### i) Disease States

5 A wide variety of disease states may be treated with compositions according to the present invention. In essence, any disease that can be treated by provision of a protein or nucleic acid is amenable to this approach. Disease states include a variety of genetic abnormalities such as diabetes, cancer, cystic fibrosis and various other diseases that could be treated by increasing or decreasing expression of a protein in a target cell.

10

Depending on the particular disease to be treated, administration of therapeutic compositions according to the present invention will be via any common route so long as the target tissue is available via that route. This includes oral, nasal, buccal, rectal, vaginal or topical. Topical administration would be particularly advantageous for treatment of skin cancers. Alternatively, administration will be by orthotopic, intradermal, subcutaneous, intramuscular, intraperitoneal or intravenous injection. Such compositions would normally be administered as pharmaceutically acceptable compositions that include physiologically acceptable carriers, buffers or other excipients.

20

20 In certain embodiments, *ex vivo* therapies also are contemplated. *Ex vivo* therapies involve the removal, from a patient, of target cells. The cells are treated outside the patient's body and then returned. One example of *ex vivo* therapy would involve a variation of autologous bone marrow transplant. Many times, ABMT fails because some cancer cells are present in the withdrawn bone marrow, and return of the bone marrow to the treated patient results in repopulation of the patient with cancer cells. In one embodiment, however, the withdrawn bone marrow cells could be treated while outside the patient with an LDL-DNA particle that targets and kills the cancer cell. Once the bone marrow cells are "purged," they can be reintroduced into the patient.

The treatments may include various "unit doses." Unit dose is defined as containing a predetermined-quantity of the therapeutic composition calculated to produce the desired responses in association with its administration, *i.e.*, the appropriate route and treatment regimen. The quantity to be administered, and the particular route and formulation, are within the skill of those in the clinical arts. Also of import is the subject to be treated, in particular, the state of the subject and the protection desired. A unit dose need not be administered as a single injection but may comprise continuous infusion over a set period of time. Unit dose of the present invention may conveniently may be described in terms of 0.01mg DNA/kg body weight to 0.4mg DNA/kg body weight, with ranges in between these being contemplated such that 0.05, 0.10, 0.15, 0.20, 0.25, 0.5mg/DNA/kg body weight are administered. Likewise the amount of LDL delivered can vary from about 0.2 to about 8.0 mg/kg body weight. Thus in particular embodiments, 0.4 mg, 0.5 mg, 0.8 mg, 1.0 mg, 1.5 mg, 2.0 mg, 2.5 mg, 3.0 mg, 4.0 mg, 5.0 mg, 5.5 mg, 6.0 mg, 6.5 mg, 7.0 mg and 7.5 mg of LDL may be delivered to an individual *in vivo*. The dosage of DNA:LDL to be administered depends to a great extent on the weight and physical condition of the subject being treated as well as the route of administration and the frequency of treatment. A pharmaceutical composition comprising the naked polynucleotide prebound to a highly purified lipoprotein fraction may be administered in amounts ranging from 1  $\mu$ g to 1mg polynucleotide to 1 $\mu$ g to 100mg protein. Thus, particular compositions may comprise 1 $\mu$ g, 5 $\mu$ g, 10 $\mu$ g, 20 $\mu$ g, 30 $\mu$ g, 40 $\mu$ g, 50 $\mu$ g, 60 $\mu$ g, 70 $\mu$ g, 80 $\mu$ g, 100 $\mu$ g, 150 $\mu$ g, 200 $\mu$ g, 250 $\mu$ g, 500 $\mu$ g, 600 $\mu$ g, 700 $\mu$ g, 800 $\mu$ g, 900 $\mu$ g or 1000 $\mu$ g polynucleotide that is bound independently to 1 $\mu$ g, 5 $\mu$ g, 10 $\mu$ g, 20 $\mu$ g, 3.0 $\mu$ g, 40 $\mu$ g 50 $\mu$ g, 60 $\mu$ g, 70 $\mu$ g, 80 $\mu$ g, 100 $\mu$ g, 150 $\mu$ g, 200 $\mu$ g, 250 $\mu$ g, 500 $\mu$ g, 600 $\mu$ g, 700 $\mu$ g, 800 $\mu$ g, 900 $\mu$ g or 1000 $\mu$ g, 1.5mg, 5 mg, 10 mg, 20mg, 30mg, 40mg, 50mg, 60 mg, 70mg, 80 mg, 90 mg or 100mg lipoprotein. Any amount of polynucleotide may be bound to any other amount of lipoprotein to achieve the pharmaceutical concentrations of the present invention.

ii) **Cancer**

One of the preferred embodiments of the present invention involves the use of the LDL vectors to deliver therapeutic genes to cancer cells. Target cancer cells include cancers of the lung, brain, prostate, kidney, liver, ovary, breast, skin, stomach, esophagus, 5 head & neck, testicles, colon, cervix, lymphatic system and blood. Of particular interest are non-small cell lung carcinomas including squamous cell carcinomas, adenocarcinomas and large cell undifferentiated carcinomas.

According to the present invention, one may treat the cancer by directly injection 10 a tumor with the LDL vector. Alternatively, the tumor may be infused or perfused with the vector using any suitable delivery vehicle. Local or regional administration, with respect to the tumor, also is contemplated. Finally, systemic administration may be performed. Continuous administration also may be applied where appropriate, for example, where a tumor is excised and the tumor bed is treated to eliminate residual, 15 microscopic disease. Delivery via syringe or catheterization is preferred. Such continuous perfusion may take place for a period from about 1-2 hours, to about 2-6 hours, to about 6-12 hours, to about 12-24 hours, to about 1-2 days, to about 1-2 weeks or longer following the initiation of treatment. Generally, the dose of the therapeutic composition via continuous perfusion will be equivalent to that given by a single or multiple 20 injections, adjusted over a period of time during which the perfusion occurs.

For tumors of  $\geq 4$  cm, the volume to be administered will be about 4-10 ml (preferably 10 ml), while for tumors of  $< 4$  cm, a volume of about 1-3 ml will be used (preferably 3 ml). Multiple injections delivered as single dose comprise about 0.1 to 25 about 0.5 ml volumes. The LDL-DNA particles may advantageously be contacted by administering multiple injections to the tumor, spaced at approximately 1 cm intervals.

In certain embodiments, the tumor being treated may not, at least initially, be 30 resectable. Treatments with therapeutic constructs may increase the resectability of the tumor due to shrinkage at the margins or by elimination of certain particularly invasive

portions. Following treatments, resection may be possible. Additional treatments subsequent to resection will serve to eliminate microscopic residual disease at the tumor site.

5           A typical course of treatment, for a primary tumor or a post-excision tumor bed, will involve multiple doses. Typical primary tumor treatment involves a 6 dose application over a two week period. The two week regimen may be repeated one, two, three, four, five, six or more times. During a course of treatment, the need to complete the planned dosings may be reevaluated.

10

Cancer therapies also include a variety of combination therapies with both chemical and radiation based treatments. Combination chemotherapies include, for example, cisplatin (CDDP), carboplatin, procarbazine, mechlorethamine, cyclophosphamide, ifosfamide, melphalan, chlorambucil, bisulfan, nitrosurea, 15 dactinomycin, daunorubicin, doxorubicin, bleomycin, plicomycin, mitomycin, etoposide (VP16), tamoxifen, taxol, transplatinum, 5-fluorouracil, vincristin, vinblastin and methotrexate.

20

Combination radiation therapies may be x- and  $\gamma$ -irradiation. Dosage ranges for x-irradiation range from daily doses of 2000 to 6000 roentgens for prolonged periods of time (3 to 4 weeks), to single doses of 2000 to 6000 roentgens. Dosages for radioisotopes vary widely, and depend on the half-life of the isotope, the strength and type of radiation emitted, and the uptake by neoplastic cells.

25

Various combinations may be employed, gene therapy is "A" and the radio- or chemotherapeutic agent is "B":

A/B/A B/A/B B/B/A A/A/B A/B/B B/A/A A/B/B/B B/A/B/B  
B/B/B/A B/B/A/B A/A/B/B A/B/A/B A/B/B/A B/B/A/A  
5 B/A/B/A B/A/A/B A/A/A/B B/A/A/A A/B/A/A A/A/B/A

The terms "contacted" and "exposed," when applied to a cell, are used herein to describe the process by which a therapeutic construct and a chemotherapeutic or radiotherapeutic agent are delivered to a target cell or are placed in direct juxtaposition with the target cell. To achieve cell killing or stasis, both agents are delivered to a cell in a combined amount effective to kill the cell or prevent it from dividing.

The therapeutic compositions of the present invention are advantageously administered in the form of injectable compositions either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection may also be prepared. These preparations also may be emulsified. A typical composition for such purpose comprises a pharmaceutically acceptable carrier. For instance, the composition may contain 10 mg, 25 mg, 50 mg or up to about 100 mg of human serum albumin per milliliter of phosphate buffered saline.

Other pharmaceutically acceptable carriers include aqueous solutions, non-toxic excipients, including salts, preservatives, buffers and the like. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oil and injectable organic esters such as ethyloleate. Aqueous carriers include water, alcoholic/aqueous solutions, saline solutions, parenteral vehicles such as sodium chloride, Ringer's dextrose, etc. Intravenous vehicles include fluid and nutrient replenishers. Preservatives include antimicrobial agents, anti-oxidants, chelating agents and inert gases. The pH and exact concentration of the various components the pharmaceutical composition are adjusted according to well known parameters.

30

Additional formulations are suitable for oral administration. Oral formulations include such typical excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate and the like. The compositions take the form of solutions, suspensions, tablets, pills, 5 capsules, sustained release formulations or powders. When the route is topical, the form may be a cream, ointment, salve or spray.

## 11. EXAMPLES

The following examples are included to demonstrate preferred embodiments of 10 the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventor to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the 15 present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

### EXAMPLE 1

#### MATERIALS AND METHODS

##### 20 1. Isolation of Plasma Lipoproteins

Restriction endonucleases were purchased from Life Technologies, and Protease 25 inhibitors (*i.e.*, leupeptin, PMSF, and Trasylol) were purchased from Sigma Chemical Company. Plasma lipoproteins were isolated using standard sequential flotation ultracentrifugation methods as described (Schumaker and Puppione, 1986). Throughout the entire procedure samples were kept on ice or at 4°C unless otherwise stated.

Subjects were fasted for at least 4 h prior to the start of the experimental 30 procedures. Blood was drawn into sterile, vacuumed glass tubes containing anticoagulants, *e.g.*, 0.1% (ethylenedinitrolo)-tetracetic acid (EDTA) or heparin. Plasma was obtained by centrifugation (10 minutes at 3000  $\times$  g) and immediately adjusted to

0.005% phenylmethansulfonyl fluoride (PMSF), 10KIU Trasylol/ml, and 1  $\mu$ g leupeptin/ml. VLDL, LDL, and HDL fractions were isolated by sequential flotation ultracentrifugation for 18 h at 40,000 rpm in a Beckmann centrifuge Model LS-80M after plasma samples were adjusted with potassium bromide (ICBr) to solution densities of 5 1.006, 1.019, and 1.215 g/ml respectively. Immediately following ultracentrifugation, individual lipoprotein fractions were collected and dialyzed extensively against phosphate buffered saline (pH 7.4) containing 0.001% sodium azide. Protein concentrations were determined using standard BCA protein assays (Pierce Chemical Company).

10 2. **Dna-Binding Protocol**

Lipoproteins and DNA were mixed together and incubated for 30 min at room temperature in 50 mmmole/liter Tris (pH 7.4), 100-154 mmoles/liter sodium chloride (NaCl), 15 mmoles/liter magnesium chloride ( $MgCl_2$ ). 6X Sample loading buffer (30% glycerol, 0.25% Xylene cyanole FF, 0.25% bromophenol blue) was added to the samples 15 in a 1:5 V/V ratio. Samples were underloaded into 30  $\mu$ l wells at the cathode edge of an 0.8% agarose gel containing 1  $\mu$ g ethidium bromide/ml in Tris-Acetate buffer (pH 7.85) and electrophoresis was accomplished using 100 Volt constant until the negatively charged tracking dye had migrated at least 50% of distance from the loading well to the anodic edge of the gel.

20

3. **Agarose Electrophoretogram of Human Lipoproteins**

Agarose electrophoresis of human lipoproteins has been performed to illustrating the differential migration patterns of lipoprotein fractions VLDL, LDL, and HDL isolated from human plasma resolved using non-denaturing conditions.

25

Plasma lipoproteins were isolated from human blood according to the protocol described above. 6X Sample loading buffer (30% glycerol, 0.25% Xylene cyanole FF, 0.25% bromophenol blue) was added to the samples in a 1:5 V/V ratio. Samples were underloaded into 30  $\mu$ l wells at the cathode edge of an 0.8% agarose gel in Tris-Acetate 30 buffer (pH 7.85) and electrophoresis was accomplished using 100 Volt constant until the

negatively charged tracking dye had migrated at least 50% of the distance from the loading well to the anodic edge of the gel.

Following electrophoresis, the agarose gel was stained for protein in a solution containing 50% V/V ethanol, 10% V/V acetic acid, and 0.25% Coomasie Brilliant Blue R-250 (CBB R-250, Bio-Rad Labs). Lane 1 contained human VLDL (10 µg protein), Lane 2 contained human LDL (35 µg protein), and Lane 3 contained human HDL (35 µg protein). Results illustrated the differential migration of lipoprotein fractions, VLDL, LDL, and HDL, isolated from human plasma resolved using non-denaturing conditions by agarose gel electrophoresis. Lipoproteins were visualized using a protein binding dye, Coomassie Brilliant Blue (CBB). The absence of other bands in each lane indicated the high degree of purity for each lipoprotein.

#### 4. Radioisotope Labeling of Deoxyoligonucleotides

Complementary single stranded oligonucleotides were mixed (10 µg each) and incubated at 85°C for 5 min in 10 mM Tris HCl (pH 7.4). Immediately following incubation, the samples were cooled down slowly to room temperature to obtain double stranded oligonucleotides. The double stranded oligonucleotides were then digested with *Bam*HI and *Eco*RI for 1 h at 37°C in 50 mM Tris HCl (pH 8.0), 100 mM NAG1, and 10 mM MgCl<sub>2</sub>. Digested double stranded oligonucleotides were purified using a Qiaquick nucleotide removal kit from Qiagen Inc. according to manufacturer's protocol. The 5' protruding ends of the purified oligonucleotides were then labeled with <sup>32</sup>P- $\alpha$ dATP using a Prime-It II labeling kit containing Exo (-) Klenow enzyme from Stratagene Inc. according to the manufacturer's protocol. The specific activity of all oligonucleotides was determined by scintillation counting.

The DNA-binding studies were performed as described above except that the agarose gel was not stained with ethidium bromide. Instead, following electrophoresis, the agarose gel was dried under vacuum and exposed to X-ray film for 4 h at room temperature prior to protein staining in a solution containing 50% V/V ethanol, 10% V/V

acetic acid, and 0.25% Coomassie Brilliant Blue R-250 (Bio-Rad Labs). Oligonucleotides and human LDL were present at 400,000 cpm and 40  $\mu$ g protein per lane respectively.

5       **5. Sonication of plasma lipoproteins**

Solutions of plasma lipoproteins in phosphate-buffered saline containing 10 mM MgCl<sub>2</sub> were kept on ice and sonicated for various time periods ranging from 0 to 6 minutes in a Sonifier Model 350 sonicator (Branson Sonic Power Co.) at the following settings: duty cycle; 30%, pulsed, output control; level 2. Immediately following 10 sonication, genomic DNA was added to the sonicated solutions, and the DNA-binding assay (see above) was started.

6.       **RT-PCR<sup>TM</sup> of Lipoprotein-bound RNA**

Human liver RNA, complexed to human LDL or to human VLDL as described 15 above, was subjected to agarose gel electrophoresis and extracted from the gel by solubilizing the gel for 20 min at 50°C in 3 times the gel volume of QX-1 buffer (Qiagen) and by twice adding an equivalent volume of phenol/chloroform (pH 4.0). RNA was precipitated by adding an equivalent volume of 100% isopropanol and freezing the mixture overnight at -80°C. RNA pellets were dissolved in 50  $\mu$ l of DEPC-treated 20 water. For each reaction, the dissolved RNA (3  $\mu$ l) was transcribed in reverse into single-stranded DNA by adding 100 mM KCl, 10 mM Tris-HCl (pH 8.3), 5 mM MgCl<sub>2</sub>, 2.5  $\mu$ M primer (oligo d(T) or random hexamers), 1 U/ $\mu$ l RNase inhibitor, 1 mM each of dATP, dCTP, dTTP, and dGTP, and 2.5 U/ $\mu$ l of MuLV reverse transcriptase in a total reaction volume of 20  $\mu$ l. The single-stranded DNA samples were then amplified in 100 mM 25 KCl, 10 mM Tris-HCl (pH 8.3), 2 mM MgCl<sub>2</sub>, 0.15  $\mu$ M each of the forward and reverse ISRE primers (see Table 2), 1 mM each of dATP, dCTP, dTTP, and dGTP, and 2.5 U/100  $\mu$ l of AmpliTaq DNA polymerase in a total reaction volume of 100  $\mu$ l. DNA amplification was carried out in a thermocycler in 30 consecutive cycles of denaturing at 95°C for 60 sec, reannealing at 55°C for 60 sec, primer extension at 72°C for 120 sec, 30 and a final extension at 72°C for 7 min. For each PCR reaction, 10  $\mu$ l of the reaction

mixture was analyzed by electrophoresis on a 1% agarose gel in TBE buffer (45 mM Tris-borate and 1 mM EDTA, pH 8.0) while maintaining a 100-V constant for 1 h. The PCR products were visualized by staining the gel with ethidium bromide.

5      **7. DNA sequencing**

DNA fragments obtained from the RT-PCR reactions were separated by electrophoresis on a 1% agarose gel and extracted from the gel by using a Qiagen gel extraction kit according to the manufacturer's protocol. DNA samples were analyzed on an Applied Biosystems Inc. model 373 automated DNA sequence apparatus after dye-  
10      terminator thermo cycle sequencing.

8.      **Cell culture and transfection assays.**

Human skin fibroblasts were cultured in complete growth medium consisting of Dulbecco's modified Eagle's medium that was supplemented with 10% fetal bovine serum, 100 µg/ml each of streptomycin and penicillin at 37°C in an atmosphere of 5% CO<sub>2</sub> in a humidified incubator. Twenty-four hours before cell transfection, during exponential growth, the cultured cells were harvested by trypsinization, replated at a cell density of  $1 \times 10^6$  cells in 35-mm culture dishes containing a glass coverslip, and cultured in complete growth medium. All transfection experiments were performed in triplicate as  
15      described.  
20

9.      **LipoFectin assay.**

The pEGFP-N1 plasmid and LipoFectin were mixed together at a ratio of 1:4 (wt/wt) in 200 µl of serum-free medium and incubated for 15 min at room temperature. When the cells reached 40 to 60% confluence, they were transfected with a mixture of 5 µg of DNA and 20 µg of LipoFectin per 35-mm culture dish, each dish having been diluted in 1 ml of serum-free medium. Transfection was performed for 16 h at 37°C. Once transfection was achieved, the liposomes were removed from the culture dish by gentle washing and maintained in 2 ml of growth medium per 35-mm culture dish for 24  
25      h at 37°C. Expression of GFP in the cells was determined by fluorescence microscopy  
30

10. **LDL assay.**

The pEGFP-N1 plasmid and LDL were mixed together at a ratio of 1:10 (wt/wt) in 100  $\mu$ l of serum-free medium containing 10 mM MgCl<sub>2</sub> and incubated for 15 min at 5 37°C. When the cells were 40 to 60% confluent, they were transfected for 16 h at 37°C with a mixture of 5  $\mu$ g of DNA and 50  $\mu$ g of LDL per 35-mm culture dish, each dish having been diluted in 1 ml of serum-free medium. Once transfection was achieved, the 10 LDLs were removed by gentle washing and maintained in 2 ml of growth medium per 35-mm culture dish for 24 h at 37°C. At 24 h after transfection, the cells were washed with PBS and fixed in 2 ml of PBS containing 4% paraformaldehyde per 35-mm culture dish for 30 min. The coverslips were then removed from the culture dishes, washed with PBS, placed in an inverted orientation on glass slides, and examined by fluorescent microscopy to detect GFP.

15. **11. *In vivo* reporter gene expression.**

Two-month-old female Sprague-Dawley rats were anesthetized with a combination anesthetic (42.8 mg/ml ketamine, 8.6 mg/ml xylazine, and 1.4 mg/ml acepromazine), and a prebound complex of purified rat LDL and linearized pEGFP-N1 plasmid DNA was injected intravenously (into the femoral vein), subcutaneously, 20 intraperitoneally, and into the pharyngeal, nasal, and rectal mucosae (100  $\mu$ g of LDL protein and 5  $\mu$ g of DNA in 100  $\mu$ l of PBS containing 10 mM MgCl<sub>2</sub> per site). Control animals were injected with linearized pEGFP-N1 plasmid DNA in which the HCMV IE promoter sequence was interrupted only by digestion with restriction enzymes. 5  $\mu$ g of DNA in 100  $\mu$ l of PBS containing 10 mM MgCl<sub>2</sub> per site. After 2, 5, or 7 days, all the 25 treated and control rats were sacrificed, their blood was collected by means of cardiac puncture, and the tissues were excised and immobilized in OCT by means of snap freezing over liquid nitrogen or by immediate freezing in liquid nitrogen. The immobilized tissue samples were sectioned on a cryomicrotome, and the sections (5–8  $\mu$ m thick) were fixed for 30 min in 4% paraformaldehyde and analyzed for expression of 30 EGFP (green fluorescent protein) by fluorescent microscopy.

**12. Fluorescent microscopy.**

Microscopy was performed by using an Olympus Model BH-2 fluorescent microscope (Olympus, USA) equipped with a digital camera (Hamamatsu, Model C5810) and a color printer (Image Master, Toshiba). The filter set used was a standard fluorescein isothiocyanate (FITC) set (Chroma Technology, Brattleboro, VT, USA). The maximum excitation and emission wavelengths for this filter set were 485 nm (range 460–510 nm) and 540 nm (range 515–565 nm), respectively. Transfection efficiency was determined by calculating the average percentage of transduced cells of five different fields per 35-mm culture dish.

**13. Detection of GFP.**

Excised rat tissues were homogenized in 150  $\mu$ l of PBS in a dounce homogenizer placed on ice. The homogenized tissues were centrifuged for 3 min at 13,000  $\times$  g, and 50- $\mu$ l aliquots were withdrawn and used in an ELISA assay to detect GFP. First, serial dilutions (range 1:10 to 1:1,000) of all samples were made in PBS. ELISA plates (96 wells) were coated with the samples (three wells/sample) by incubating the plates at room temperature for 3 h. The plated samples were then washed three times with 200  $\mu$ l of 1  $\times$  PBS containing 0.1% Tween 20 (PBST) and blocked with 200  $\mu$ l of PBST containing 1% bovine serum albumin (BSA) for 2 h at room temperature while shaking gently. The washing procedure was repeated with 200  $\mu$ l of PBST containing 0.1% BSA, and the plated samples were incubated with a 1:2,000 dilution of a recombinant GFP polyclonal antibody (IgG fraction, Clontech Inc., Palo Alto, CA) in PBST containing 0.1% BSA (50  $\mu$ l of diluted mixture per well) for 18 h at 4°C while shaking gently. The plated samples were washed and incubated with a 1:5000 dilution of HRP-conjugated goat anti-rabbit antibody (IgG fraction, Cappel, Durham, NC) in PBST containing 0.1% BSA for 1 h at room temperature while shaking gently. The washing procedure was repeated and was followed by a final wash with 1  $\times$  PBS. GFP was detected after a 30-min incubation at room temperature in PBS containing  $\sigma$ -phenylenediamine as a chromogenic substrate.

## EXAMPLE 2

### BINDING OF HUMAN GENOMIC DNA TO HUMAN LDL

5 The binding of human genomic DNA (hg DNA) to human LDL has also been demonstrated. Each lane of the agarose gel contained hg DNA cut with *AluI* or *HindIII*. In addition, human VLDL and mouse LDL were run alongside the hg DNA.

10 Plasma lipoproteins were isolated from human or mouse blood according to the protocol described above. DNA-binding studies were performed using human genomic DNA digested with either *AluI* or *HindIII*. Following electrophoresis, the gel was stained for DNA with ethidium bromide prior to protein staining in a solution containing 50% V/V ethanol, 10% V/V acetic acid, and 0.25% Coomasie Brilliant Blue R-250 (CBB R-250, Bio-Rad Labs).

15 Each lane contained 5 µg human genomic DNA (hg DNA) cut with *AluI* or *HindIII*. In addition, human VLDL (10 µg protein per lane) human LDL (35 µg protein per lane) and mouse LDL (10 µg protein per lane) were also analysed.

20 Bands in this study showed specific binding of digested human DNA fragments and human LDL by gel-shift electrophoresis. DNA fragment obtained by *AluI* or *HindIII* digestion of human genomic DNA are shown to migrate toward the anode with much slower mobility when preincubated with human LDL but not when incubated with human VLDL, human HDL, or mouse LDL. The complexed DNA/lipoprotein band are first visualized using DNA-binding ethidium bromide and photographed using transmitted 25 ultra-violet light for activation of the fluorescent dye. Lipoproteins were next visualized with CBB and photographed using transmitted visible light. The results shown in this figure indicate that aliquots of *AluI*- and *Hind* III-digested human genomic DNA fragments comigrate with human LDL and are therefore bound to human LDL.

While *AluI*, and *HindIII* were used to digest genomic DNA in the studies shown here, the inventors of the instant invention have also used *BamHI*, and *PvuI* for genomic DNA digest. It is understood by those of skill in the art that there are many known restriction enzymes. All of which are capable of genomic DNA digestion resulting in 5 DNA that can be successfully bound to LDL. DNA digested with *AluI* yields DNA of very small size (200-700 nucleotides) which allows isolation of the slower migrating digested DNA bound to LDL from the unbound digested DNA using agarose gel electrophoresis. Digestion of genomic DNA with *HindIII* yields genomic DNA of greater average size (1000-7000 nucleotides) which reaches the upper size limit for separation by 10 agarose gel electrophoresis (the technique used here), however there are other known DNA separation techniques which would work similarly to accomplish the goal of separating free DNA from DNA bound to LDL. The choice of which separation technique to use is dependent only on the size of the DNA fragments resulting after digestion. In principal, undigested genomic DNA would also work.

15

### **EXAMPLE 3**

#### **BINDING OF PLASMID DNA TO HUMAN LDL**

Plasma LDL were isolated from human blood according to the protocol previously described in Example 1. DNA-binding studies were using DNA (pBluescript 20 II KS, Stratagene Inc.) digested with *Pvu* I. Following electrophoresis, the agarose gel was stained for DNA with ethidium bromide prior to protein staining in a solution containing 50% V/V ethanol, 10% V/V acetic acid, and 0.25% Coomassie Brilliant Blue R-250 (CBB R-250, Bio-Rad Labs). The binding of plasmid DNA to human LDL was shown in agel which contained contains 0.5 µg molecular size DNA marker (Lane 1); 2 25 µg pKS DNA cut with *Pvu* I (Lanes 2-4); 35 µg human LDL (Lane 3) and 70 µg human LDL protein (Lane 4).

Results of the electrophoretogram illustrated specific binding of *PvuI* digested plasmid DNA (pBluescript II KS, Stratagene Inc.) and human LDL. Increased amounts

of human LDL also caused an increase of DNA shifted to the LDL location and a decrease of the free *Pvu* I digested DNA band. Co-migration of the *Pvu* I digested DNA and human LDL are proof of a physical complex composed of LDL and DNA.

5

#### EXAMPLE 4

#### BINDING OF CMV PROMOTER-REGULATORY

#### SEQUENCES TO HUMAN LDL

Plasma lipoproteins were isolated from human or mouse blood according to the protocol previously described in Example 1. DNA-binding studies were performed using 10 plasmid DNA (either pBluescript II KS or pBKCMV, Stratagene Inc.) digested with *Bam*HI. Following electrophoresis the agarose gel was stained for DNA with ethidium bromide prior to protein staining in a solution containing 50% V/V ethanol, 10% V/V acetic acid, and 0.25% Coomassie Brilliant Blue R-250 (CBB R-250, Bio-Rad Labs). Loading quantities per lane were as follows:

15 plasmid DNA: 1 µg DNA/lane  
human VLDL 35 µg protein/lane  
human LDL 35 µg protein/lane  
mouse VLDL: 8 µg protein/lane  
mouse LDL: 35 µg protein/lane

20 This study used *Bam*HI cut pIGS, *Bam*HI cut pBKCMV, human VLDL, human LDL, mouse VLDL and mouse LDL.

A comparison of human LDL complexed with *Bam*HI linearized plasmids, pBluescript II KS or pBKCMV. The inventors' results illustrated that specific binding of 25 *Bam*HI linearized plasmid DNA and human LDL occurs, but these *Bam*HI linearized plasmids do not complex with either human VLDL, mouse VLDL or mouse LDL under the conditions previously described in the DNA-binding protocol (Example 2). Further, enhanced binding of human LDL and the *Bam*HI linearized plasmid pBKCMV DNA which contains the cytomegalovirus promoter region SEQ ID NO:225 (Table 2) was 30 observed as compared to the *Bam*HI linearized plasmid pBluescript II KS DNA that does

not contain the cytomegalovirus promoter region (lane 3). Because binding of DNA by LDL is enhanced in the presence of the CMV promoter, it is possible that 'LDL binds specifically to the CMV promoter sequence (SEQ ID NO:225, see Table 2).

5           Aliquots containing approximately 8 µg mouse VLDL protein were used in each DNA-binding assay mixtures resolved in lanes 4 and 9 as compared to 35 µg of total protein of all other lipoproteins (lanes 2, 3, 5, 7, 8, and 10). Due to the low physiological concentration of VLDL in mouse plasma and the limited loading capacity of the gel, it was not possible to load 35 µg of mouse VLDL protein per lane. Therefore, this study  
10          does not allow for a quantitative comparison of the plasmid DNA-binding capacity of mouse VLDL *vs.* human VLDL, human LDL, and mouse LDL.

**TABLE 2**  
**Nucleotide Sequence of the Promoter Region (1300-1900) of the Human**  
**Cytomegalovirus SEQ ID NO:225**

GGATCTGACG	GTTCACTAAA	CCAGCTCTGC	TTATATAGAC	CTCCCACCGT
ACACGCCTAC	<u>CGCCCATTG</u>	<u>CGTCAATGGG</u>	<u>GCGGAGTTGT</u>	TACGACATTT
TGGAAAGTCC	<u>CGTTGATT</u> TT	<u>GGTGCCAAA</u> AA	<u>CAAAC</u> TCCAT	<u>TGACGTCA</u> AT
<u>GGGGTGGAGA</u>	CTTGGAAATC	CCCGTGAGTC	AAACCGCTAT	CCACGCCAT
<u>TGATGTACTG</u>	<u>CCAAAACCGC</u>	ATCACCATGG	TAATAGCGAT	GACTAATACG
<u>TAGATGTACT</u>	<u>GCCAAGTAGG</u>	AAAGTCCCAT	AAGGTCAATGT	ACTGGGCATA
ATGCCAGGCG	<u>GGCCATT</u> TAC	<u>CGTCATTGAC</u>	<u>GTCAATAGGG</u>	<u>GGCGTACT</u> TG
GCATATGATA	CACTTGATGT	ACTGCCAAGT	GGGCAGTTA	CCGTAAATAC
TCCACCCATT	<u>GACGTCAATG</u>	<u>GAAAGTCCCT</u>	ATTGGCGTTA	CTATGGGAAC
<u>ATACGTCA</u> TT	<u>ATTGACGTCA</u>	<u>ATGGGCGGGG</u>	<u>GTCGTTGGC</u>	GGTCAGCCAG
GC <del>GGG</del> CCATT	<u>TACCGTAAGT</u>	TATGTAACGC	GGAACTCCAT	ATATGGGCTA
TGA <del>ACT</del> TAATG	<u>ACCCCGTAAT</u>	TGATTACTAT	TAATAACTA	

Major repeat regions are indicate in bold and underlined.

5

**EXAMPLE 5**  
**BINDING OF SRE, E/C, FAS, AND ISRE**  
**DEOXYNUCLEOTIDE SEQUENCES TO HUMAN LDL**

10 Plasma lipoproteins were isolated from human or mouse blood according to the protocol previously described in Example 1. DNA-binding studies were performed using the synthetic oligonucleotides: SRE, E/C, and FAS (see Table 3 for nucleotide sequences).

**TABLE 3**  
**Deoxyribonucleic Acid Sequences of Synthetic Oligonucleotides**  
**used in Binding Studies with LDL**

SEQ ID NO	Oligo Name	Sequence (5'-3')
226	SRE-2A	GATCCAAATCACCCACTGCAACTCCTCCCCCTGCG
227	E/C-1A	GATCCATCCAATTGGGCAATCAGGAG
228	FAS- 1A	GATCCGGTCTCCAATTGG
229	ISRE- 1A	GATCCTCGGGAAAGGGAAACCGAAACTGAAGCCG

5 DNA-binding studies were performed according to the previously described DNA-binding protocol (Example 2). Following electrophoresis, the agarose gel was stained for DNA with ethidium bromide prior to protein staining in a solution containing 50% V/V ethanol, 1096 V/V acetic acid, and 0.25% Coomassie Brilliant Blue R-250 (CBB R-250, Bio-Rad Labs). Oligonucleotides were present at 1 µg DNA per lane.  
10 Lanes containing human LDL contained 35 µg protein per lane and lanes containing mouse LDL contained 15 µg protein per lane.

15 The data generated showed the complexed synthetic, double-stranded oligonucleotide fragments and human LDL. The results strongly support that human LDL binds to these DNA sequences in a highly specific fashion. The synthetic oligonucleotides SRE-2A, E/C-1A, FAS-1A, and ISRE-1A (Table 3, SEQ ID NO:226, SEQ ID NO:227, SEQ ID NO:228, and SEQ ID NO:229 respectively) bind to human LDL but do not bind to mouse LDL. DNA binding to human LDL is illustrated by the appearance of a fraction of slower mobility DNA that comigrates with human LDL.

20 In another embodiment of this same study, binding was determined using radioisotope labeling of the deoxynucleotide sequences as described in Example 1. The results from these DNA-binding studies show that human LDL binds to the synthetic oligonucleotides SRE-2A, E/C-1A, FAS-1A, and ISRE-1A (Table 3, SEQ ID NO:226;

SEQ ID NO:227; SEQ ID NO:228; SEQ ID NO:229) in a highly specific fashion. DNA binding to human LDL is illustrated by the appearance of a fraction of slower mobility DNA that comigrates with human LDL. The binding affinity of the different synthetic oligonucleotides for human LDL can be determined by kinetic binding studies using quantitative autoradiography well known to those of skill in the art.

5

## **EXAMPLE 6**

### **BINDING OF VARIOUS NUCLEOTIDE SEQUENCES TO THE LDL ISOLATED FROM VARIOUS SPECIES**

10 Plasma lipoproteins were isolated from human, mouse, rat, or baboon blood according to the protocol previously described in Example 1. DNA-binding studies were performed according to the previously described DNA-binding protocol using the synthetic oligonucleotides: SRE, E/C, and FAS (see Table 3 for nucleotide sequences), genomic DNA, or plasmid DNA containing the CMV promoter. A summary of the 15 binding studies of the instant invention are illustrated in Tables 4A and 4B, below. Table 4A illustrates the binding of human, mouse, rat and baboon LDL to various forms and sources of DNA, and Table 4B illustrates the DNA/LDL complexes made thus far.

**TABLE 4A**  
**Binding of Human, Mouse, Rat and Baboon LDL to Various Forms of DNA**

DNA	human LDL	mouse LDL	rat LDL	baboon LDL
hg DNA	YES	NO	YES	YES
mg DNA	N.D.	N.D.	YES	N.D.
rg DNA	N.D.	N.D.	YES	N.D.
bg DNA	N.D.	N.D.	N.D.	YES
CMV	YES	NO	YES	YES
SRE	YES	NO	N.D.	NO
E/C	YES	NO	N.D.	NO
FAS	YES	NO	N.D.	NO

hg = human genomic DNA (digested with either *Alu*I or *Hind*III, mg = mouse genomic DNA digested with either *Alu*I or *Hind*III, rg = rat genomic DNA digested with either *Alu*I or *Hind*III, and bg = baboon genomic DNA digested with either *Alu*I or *Hind*III

Yes = binding, NO = no binding, N.D. = binding not determined

**TABLE 4B**  
**Specific LDL/DNA Complexes That Have Been Made**

<b>DNA</b>	<b>DNA Digested With</b>	<b>LDL</b>
human genomic	<i>AluI</i>	human
human genomic	<i>HindIII</i>	human
human genomic	<i>Bam</i> HI	human
human genomic	<i>Pvu</i> I	human
human genomic	<i>AluI</i>	rat
human genomic	<i>HindIII</i>	rat
human genomic	<i>Bam</i> HI	rat
human genomic	<i>Pvu</i> I	rat
human genomic	<i>AluI</i>	baboon
human genomic	<i>HindIII</i>	baboon
human genomic	<i>Bam</i> HI	baboon
human genomic	<i>Pvu</i> I	baboon
mouse genomic	<i>AluI</i>	rat
mouse genomic	<i>HindIII</i>	rat
rat genomic	<i>AluI</i>	rat
rat genomic	<i>HindIII</i>	rat
baboon genomic	<i>AluI</i>	baboon
baboon genomic	<i>HindIII</i>	baboon
pBSKS	<i>Pvu</i> I	human
pBSKS	<i>Bam</i> HI	human
pBKCMV	<i>Bam</i> HI	human
pBKCMV	<i>Bam</i> HI	rat
pBKCMV	<i>Bam</i> HI	baboon
SRE-2A oligo	none	human
SEQ ID NO:226		
E/C-1A oligo	none	human
SEQ ID NO:227		
FAS-1A oligo	none	human
SEQ ID NO:228		
ISRE-1A oligo	none	human
SEQ ID NO:229		

**EXAMPLE 7**  
**DETECTION OF LDL-BOUND DNA IN HUMAN BLOOD**

Plasma lipoproteins are isolated from human blood according to the protocol  
5 previously described in Example 1. 6X Sample loading buffer (30% glycerol, 0.25%  
Xylene cyanole FF, 0.25% bromophenol blue) is added to the samples in a 1:5 V/V ratio.  
Samples are underloaded into 30  $\mu$ l wells at the cathode edge of an 0.8% agarose gel in  
10 Tris-Acetate buffer (pH 7.85) and electrophoresis is accomplished using 100 Volt  
constant until the negatively charged tracking dye migrates at least 50% of the distance  
from the loading well to the anodic edge of the gel. Following electrophoresis, is stained  
for DNA with ethidium bromide prior to protein staining in a solution containing 50%  
15 V/V ethanol, 10% V/V acetic add, and 0.25% Coomasie Brilliant Blue R-250 (CBB R-  
250, Bio-Rad Labs). If no DNA is detected by ethidium bromide staining, the agarose  
gel is subjected to Southern blot analysis using a labeled DNA probe. The DNA is  
labeled with a radioactive isotope (e.g.,  $^{32}$ P), a non-radioactive tag (DIG) or with any  
other standard DNA-labeling method known to one of skill in the art. Randomly  
synthesized, short oligonucleotides are used as the probe to detect, in a general fashion,  
whether or not DNA is bound to the isolated LDL. Controls include lanes containing  
known quantities of DNA, lanes containing purified LDL digested with DNase I, and  
20 LDL bound to DNA made by mixing purified LDL and DNA according to the method  
described in Example 2.

LDL isolated from humans with cancer and subjected to the above protocol will  
have detectable DNA bound to the LDL in quantities greater than the amount of DNA  
bound to LDL isolated from humans without cancer.

**EXAMPLE 8**

**DETECTION OF SPECIFIC TYPES OF CANCERS WITH**

**SEQUENCE SPECIFIC DNA PROBES**

Not only is it possible to identify the presence or absence of cancer in a living body using the invention technique (as described in Example 14 above), it is also possible to identify specific cancer types by using sequence specific DNA probes. For example, LDL-bound DNA isolated from a patient with colon cancer will have a different DNA sequence than the LDL-bound DNA isolated from a patient with a different cancer type, for example, breast cancer. Different DNA sequences bound to the LDL isolated from different cancer patients is determined by first isolating LDL from the blood of a person with an independently identified and known cancer type, using the protocol in Example 1. This isolated LDL is then digested with various non-specific proteases to remove the LDL while retaining the DNA. This DNA is then sequenced using standard sequencing techniques. A list of the DNA sequences along with the type of cancer it is associated with is made. This list is then used to synthesize probes that can differentiate among the various types of cancer. These probes are used in screening of a patient with an unknown cancer type, or in the early detection of metastatic cancer, or as a general early screening technique for the presence or absence of specific cancer types.

20

**EXAMPLE 9**

**METHODS FOR THE DETERMINATION OF METASTATIC GENE**

**TRANSFER VIA LIPOPROTEINS AS NATIVE VECTORS**

In order to determine the sequence of polynucleotides bound to endogenous LDL, plasma LDL and other apoB-containing lipoproteins are captured using a monoclonal antibody to a specific apoB epitope such as 2G8 which is immobilized on an inert, hydrophilic and highly porous polymer microbead. The LDL-DNA complex is then isolated by elution using affinity chromatography technology. DNA is further purified from the isolated LDL/DNA complex using standard DNA purification methodology

such as phenol/chloroform extraction followed by ethanol precipitation. Alternatively, purified DNA is isolated from the affinity column using elution conditions that disrupt protein/DNA complexes but not protein/protein complexes (*i.e.*, antibody/LDL complex).  
5 The polynucleotide sequences are determined using the SRE, E/C, FAS, and ISRE-1A oligonucleotides (SEQ ID NO:226, SEQ ID NO:227, SEQ ID NO:228, and SEQ ID NO:229, respectively) in a standard PCR™ methodology in order to amplify polynucleotides with unknown sequences. The amplified PCR™ products (*i.e.*, polynucleotides) are then isolated by agarose gel electrophoresis and subsequent DNA sequencing techniques well known to the art.

10

Alternatively, identification of polynucleotide sequences that are bound to endogenous human LDL is *via* the specific binding of LDL to a plastic matrix such as a 96 well ELISA (enzyme linked immunosorbant assay) plates coated with specific antibodies that bind to human LDL. In this embodiment, freshly isolated plasma containing endogenous lipoproteins is used to bind to the anti-human LDL antibodies using standard ELISA procedures lipoproteins to the art. The presence and specific sequence of polynucleotides prebound to the endogenous LDL in each is determined by PCR™ technology.  
15

20

Because many varying and different embodiments may be made within the scope of the inventive concept herein taught, and because many modifications may be made in the embodiments herein detailed in accordance with the descriptive requirement of the law, it is to be understood that the details herein are to be interpreted as illustrative and not in a limiting sense.

25

**EXAMPLE 10**  
**LOW-DENSITY LIPOPROTEIN INTERACTS WITH**  
**HUMAN CYTOMEGALOVIRUS GENOMIC DNA**

5 DNA binding experiments with purified plasma lipoprotein fractions and human  
genomic DNA as well as several different plasmids indicate that purified LDL binds to  
human genomic DNA digested with different restriction enzymes (Alu I and Hind III).

Purified LDL also bound to several different plasmids but its binding affinity for  
plasmid DNA containing the HCMV IE promotor region was significantly higher. It was  
10 shown that the binding of both LDL and VLDL to the HCMV IE promotor region and  
SRE, MSRE, ISRE, MISRE, E/C, FAS, and MFAS oligonucleotides. The E/C  
oligonucleotide was used in these DNA binding studies because this oligonucleotide  
contains both a binding site for members of the C/EBP transcription factor family, which  
are involved in the regulation of differentiation-dependent adipocyte gene expression, as  
15 well as an overlapping E-box motif which is generally recognized by the eukaryotic basic  
helix-loop-helix (b-HLH) transcriptional regulators. LDL clearly have a greater affinity  
for all of the oligonucleotides tested than do VLDL. This is most likely due to  
interference with protein-DNA interaction caused by either the presence of other  
apolipoproteins on the surface of VLDL or an increased net charge as a result of the  
20 increased lipid content of VLDL.

The sequence specificity is illustrated by the fact that both LDL and VLDL show  
a decreased binding affinity for the mutated versions of the ISRE and FAS oligos  
(MISRE and MFAS respectively). In contrast, LDL showed an increased binding affinity  
25 for the mutated version of the SRE oligo (MSRE). It is possible that this mutated SRE  
sequence may be a better ligand for the putative DNA binding region of apo B present on  
LDL. The binding of both VLDL and LDL to the E/C oligonucleotide is not surprising  
since this oligo contains the E-box motif which is a known binding site for b-HLH  
proteins and similar b-HLH regions have been identified in apoB present on VLDL and  
30 LDL.

The affinity for the HCMV IE promotor is not immediately obvious since careful analysis does not reveal an exact copy of either a SRE, ISRE, FAS, or E/C sequence. However, the HCMV IE promotor region contains regulatory elements that are generally 5 recognized by a large number of eukaryotic DNA-binding proteins, including a variety of different families of transcription factors, and it may therefore be possible that the identified b-HLH regions of apoB possess similar DNA binding properties.

Another possibility is that other yet unidentified regions of apoB are involved in 10 the binding to the HCMV IE promotor region. The fact that HDL in contrast to VLDL and LDL do not bind to any of the oligos tested suggests that the DNA binding results from the specific interaction with apo B. These data support the hypothesis that apo B contains DNA binding domains which show homology with the DNA binding domains of 15 SREBP-1, SREBP-2, ADD-1, and ISGF3 $\gamma$  and that apo B containing lipoproteins therefore bind to specific nucleotide sequences similar to those bound by these known DNA binding proteins.

Recent reports suggest a possible causal relationship between human 20 cytomegalovirus (HCMV) and the development of atherosclerosis in humans. These reports together with data presented herein, which show that human LDL binds strongly to HCMV IE promotor sequences, led the inventors to investigate whether plasma LDL may play a role in the pathogenesis of HCMV induced atherosclerosis.

To test this hypothesis, the inventors looked for HCMV DNA sequences in the 25 purified plasma LDL fraction of human subjects who tested seropositive for HCMV by polymerase chain reaction (PCR). The results of these studies show that a PCR product of the expected size (170 bp) could be detected with both primer sets (MTR2 and IE) in the purified plasma LDL fraction of HCMV seropositive subjects. However, this 170 bp DNA fragment could not be detected in the plasma samples of these subjects (lanes 6-8). 30 These data suggest that the use of purified plasma LDL fractions for detection of CMV

nucleic acid sequences by PCR techniques is more sensitive than when whole plasma samples are used. Furthermore, the increased yield of PCR products of the purified plasma LDL fractions strongly suggest that HCMV DNA is predominantly associated with LDL within the plasma pool of HCMV seropositive subjects.

5

**EXAMPLE 11**  
**LOW-DENSITY LIPOPROTEIN AS A**  
**NATURAL GENE TRANSFER VECTOR**

The discovery of the nucleic acid-binding properties apo B-100 suggested that  
10 lipoproteins containing apoB100, as naturally occurring liposomes, may function as gene transfer agents. By using highly purified low-density lipoprotein as such an agent, the inventors were able to transfect cultured human skin fibroblasts *in vitro* and to express a green fluorescent protein reporter gene *in vivo*. The gene transfer mediated by low-density lipoprotein was more efficient than that mediated by LipoFectin. Low-density  
15 lipoprotein also did not exhibit any toxicity, immunogenicity, or serum inhibition.

**1. DNA-binding**

In the Examples above, it was shown that highly purified human LDL binds to  
nucleic acids in a specific fashion. In order to establish whether rat lipoproteins can bind  
20 nucleic acids in a similar fashion, DNA-binding experiments with different rat lipoprotein fractions were performed. A gel shift assay of linearized pBluescript KS and pBKCMV plasmid DNA and purified rat VLDL, LDL, and HDL fractions was performed. The data clearly demonstrate that the binding of nucleic acids is specific to the purified LDL fraction.

25

The binding of LDL to DNA is exhibited by the retarded electrophoretic migration of DNA in agarose gel that is caused by the formation of complexes of higher molecular weight. In contrast, purified fractions of VLDL and HDL did not bind any of the DNA samples tested. The fact that purified HDL did not bind DNA was expected,  
30 since endogenous HDL does not contain apo B-100. Surprisingly, there was no apparent

binding of DNA to apo B-100-containing VLDL. It is possible that the DNA-binding assay, which employs ethidium bromide staining to detect DNA, lacks sensitivity or that VLDL does not bind to DNA under the conditions of the DNA-binding assay. Another explanation could be a difference in the conformation of apo B-100 present on LDL as opposed to VLDL because of a difference in the lipid composition and protein content of the two lipoprotein fractions.

5           2.     *In vitro* cell transfection studies.

Based on the findings of the DNA-binding assay, transfection studies were 10 performed using a prebound complex of LDL and plasmid DNA that contained a reporter gene that encodes GFP.

15           The data generated illustrated the successful transfection of how human skin fibroblasts with LDL and pEGFP-N1 plasmid DNA. The transfection process was monitored by expression of the GFP encoding gene and is driven by the HCMV IE promoter. In addition to fluorescent microscopic analysis, expression of GFP was confirmed by a qualitative ELISA using a primary antibody against recombinant GFP and an HRP-conjugated secondary antibody with  $\sigma$ -phenylenediamine as a chromogenic substrate.

20

Human skin fibroblasts transfected with LDL exhibited a significantly lower 25 intensity of green fluorescence than did cells transfected with LipoFectin, indicating that the level of GFP expression was lower in these LDL-transfected cells. When the percentage of positively transfected cells were compared, however, transfection with LDL yielded a higher percentage of transfected cells than did transfection with LipoFectin (20 to 30% and 60 to 70%, respectively). In addition, LipoFectin-mediated transfection resulted in green fluorescence in the cell cytoplasm and in the nuclei, whereas LDL-mediated transfection resulted in green fluorescence predominantly in the cytoplasm.

Transfection assays in which LDL concentrations were as high as 250 g/ml of LDL protein produced no detectable effects on the confluence and viability of the cell cultures, whereas LipoFectin concentrations of 20 g/ml resulted in significant loss of cell viability. Control cells that were transfected with linearized pEGFP-N1 plasmid DNA only exhibited no fluorescence.

5           3. *In vivo reporter gene expression.*

To evaluate whether LDL could be used as a vehicle for *in vivo* gene delivery, a prebound rat LDL-pEGFP-N1 complex was administered to 2-month-old female Sprague-Dawley rats. Cryosections of the liver and heart tissues of the treated animals that had been excised 2 days after the LDL-pEGFP-N1 complex showed significant levels of green fluorescence indicative of EGFP expression as determined by fluorescent microscopy.

15           The expression of GFP in the different tissues was confirmed by a qualitative ELISA using a primary antibody against recombinant GFP and an HRP-conjugated secondary antibody with  $\sigma$ -phenylenediamine as a chromogenic substrate. In contrast, only low levels of autofluorescence were observed in the cryosectioned tissues obtained 20 from the control animals treated solely with linearized pEGFP-N1 DNA. These data demonstrate that purified LDL can be used in a prebound complex with DNA as an *in vivo* gene delivery system.

\*           \*           \*

25           All of the compositions and/or methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied 30 to the compositions and/or methods and in the steps or in the sequence of steps of the

method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims

5

## REFERENCES

- U.S. Patent No. 4,396,601
- U.S. Patent No. 4,403,035
- U.S. Patent No. 4,497,796
- 5 U.S. Patent No. 4,663,292
- U.S. Patent No. 4,868,116
- U.S. Patent No. 4,885,248
- U.S. Patent No. 4,904,582
- U.S. Patent No. 5,023,243
- 10 U.S. Patent No. 5,096,815
- U.S. Patent No. 5,149,782
- U.S. Patent No. 5,168,062
- U.S. Patent No. 5,198,346
- U.S. Patent No. 5,219,740
- 15 U.S. Patent No. 5,252,479
- U.S. Patent No. 5,283,185
- U.S. Patent No. 5,298,422
- U.S. Patent No. 5,385,839
- U.S. Patent No. 5,482,853
- 20 U.S. Patent No. 5,518,913
- U.S. Patent No. 5,521,291
- U.S. Patent No. 5,523,222
- U.S. Patent No. 5,547,932
- U.S. Patent No. 5,574,142
- 25 U.S. Patent No. 5,578,475
- U.S. Patent No. 5,580,558
- U.S. Patent No. 5,580,859
- U.S. Patent No. 5,585,254

U.S. Patent No. 5,585,362

U.S. Patent No. 5,591,601

Baleja and Sykes, *Biochem. Cell Biol.*, 69(2-3):202-205, 1991.

5 Banner, Kolddnidis, Tsemoglou, *J. Mol. Biol.*, 196(3):657-675, 1987.

Benvenisty *et al.*, *Proc. Natl. Acad. Sci. USA*, 83:9551-9555, 1986.

Biou, Yaremchuk, Tukaio, Cusack, *Science*, 263(5152):1404-1410, 1994.

Blanco-Vaca *et al.*, *Chem. Physics, Lipids*, 67/68:35-42, 1994.

Blanco-Vaca *et al.*, *J. Lipid Res.*, 33:1785-17986, 1992.

10 Briggs *et al.*, *J. Biol. Chem.*, 268:14490-14496, 1993.

Brown and Goldstein, *Science*, 232:34-47, 1986.

Bullough, Hughson, Skehel, Wiley, *Nature*, 371(6492):37-43, 1994b.

Bullough, Hughson, Treharne, Ruigrok, Skehel, Wiley, *J. Mol. Biol.*, 236(4):1262-1265, 1994a.

15 Cardin *et al.*, *Biochem.*, 26:5513-5518, 1987.

Carpenter, *FASEB J.*, 6:3283-3289, 1992.

Cicchetti *et al.*, *Science*, 257:803-806, 1992.

Cohen and Parry, *TIBS*, 11:245-248, 1986.

Coleman *et al.*, *Biochim. Biophys. Acta*, 1037:129-132, 1990.

20 Connelly *et al.*, *J. Lipid Res.*, 34:1717-1727, 1993.

Cooper, "The SRC family of protein-tyrosine kinases," Chapter 3, *In: Peptides and Protein Phosphorylation*, B.E. Kemp, ed., CRC Press, Inc., Boca Raton, Fl., pp. 85-113

Curiel *et al.*, *Human Gene Therapy*, 3:147-154, 1992.

25 Cusack, Berhet-Colominas, Hartlein, Nassar, Leberman, *Nature*, 347(6290):249-255, 1990.

Ellenberger, Brandl, Struhl, Harrison, *Cell*, 71(7):1223-37, 1992.

Escobedo *et al.*, *Mol. Cell Biol.*, 11: 1125-1132, 1991.

Felgner *et al.*, *Proc. Natl. Acad. Sci. USA*, 84:7413-7417, 1987.

Geahlen and Harrison, "Protein-tyrosine kinases," Chapter 10, *In: Peptides and Protein Phosphorylation*, B.E. Kemp, ed., CRC Press, Inc., Boca Raton, Fl., pp. 239-253.

Glover and Harrison, *Nature*, 373(6511):257-261, 1995.

Guevara, Jr. *et al.*, *Arterioscler.*, 13:758-770, 1993.

5 Guevara, Jr. *et al.*, *Proteins*, 12:188-199, 1992.

Guevara, Jr. *et al.*, Src signal sequences in apoprotein B-100: "Does LDL or Lp[a] play a role in signal transduction," *In: Multiple Risk Factors in Cardiovascular Disease*, Yamamoto, ed., Churchill Livingston Japan, ICIC, 1994.

Hunter and Cooper, *Ann. Rev Biochem*, 54:897-930, 1985.

10 Innerarity *et al.*, *Acta Med. Scand. Suppl.*, 715:51-59, 1986.

Kinoshita *et al.*, *J. Lipid Res.*, 31:701-708, 1990.

Knott *et al.*, *Nature*, 323:734-738, 1986.

Koch *et al.*, *Science*, 252:668-674, 1991.

Kostner and Grillhofer, *J. Biol. Chem.*, 266:21287-21292, 1991.

15 Kraulis, Raine, Gadhixji, Laue, *Nature*, 356(6368):448-450, 1992.

Lee *et al.*, *Human Gene Therapy*, 7:1701-1717, 1996.

Lowenstein *et al.*, *Cell*, 70:431-442, 1992.

Lupas, Van Dyke, Stock, *Science*, 252(5010):1162-1164, 1991.

MacArthur and Thornton, *J. Mol. Biol.*, 218:397-412, 1991.

20 Mahley, *Science*, 240:622-630, 1988.

Marcel *et al.*, *Arterioscl.*, 8:832, 1988.

Margolis, *Cell, Growth and Differentiation*, 3:73-80, 1992

Marmorstein, Carey, Ptashne, Harrison, *Nature*, 356(6368):408-414, 1992.

McLean *et al.*, *Nature*, 330:132-137, 1987.

25 Milne *et al.*, *J. Biol. Chem.*, 264(33):19754-19760, 1989.

Musacchio *et al.*, *Nature*, 359:851-855, 1992.

Myant, "LDL: Physical and Chemical Characteristics, Chapter 5, *In: Cholesterol Metabolism, LDL, and the LDL Receptor*, San Diego, Academic Press, pp. 99-111, 1990.

30 Neumann *et al.*, *EMBO J.*, 7:841-845, 1982.

Noble *et al.*, *EMBO J.*, 12:2617-2624, 1993.

O'Shea, Klemm, Kim, Alber, *Science*, 254(5031):539-544, 1991.

Pawson, *Current Opinion in Structural Biology*, 2:432-437, 1992.

Pleiman *et al.*, *Science*, 263:1609-1612, 1994.

5 Ren *et al.*, *Science*, 259:1157-1161, 1993.

Rudd *et al.*, *Biochim. Biophys. Acta.*, 1155:239-266, 1993.

Schimotohono *et al.*, *Cell*, 26:67-77, 1981.

Schlessinger, *Current Opinions in Genetics and Development*, 4:25-30, 1994.

Schumaker and Puppione, *Methods in Enzymology*, 128:155-170, 1986.

10 Smith *et al.*, *J. Biol. Chem.*, 265:2306-2310, 1990.

Sompayrac *et al.*, *Proc. Natl. Acad. Sci. USA*, 78:7575-7578, 1981.

Songyang *et al.*, *Cell*, 72:767-778, 1993.

Stribling *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:11277-11281, 1992.

Teng *et al.*, *J. Biol. Chem.*, 260:5067-5072, 1985.

15 Trieu and McConathy, *Biochem.*, 29(25):5919-5924, 1990.

Trieu *et al.*, *J. Biol. Chem.*, 266(9):5480-5485, 1991.

Ugi *et al.*, *FEBS Lett.*, 340:216-220, 1994.

Veals *et al.*, *Mol. Cell. Biol.*, 12:3315-3324, 1992.

Waksman *et al.*, *Cell*, 72:779-790, 1993.

20 Wang *et al.*, *J. Biol. Chem.*, 268:14497-14504, 1993.

Weisgraber and Rall, Jr., *J. Biol. Chem.*, 262:11097-11103, 1987.

Wigler *et al.*, *Proc. Natl. Acad. Sci. USA*, 76:1373-1376, 1979.

Wolff *et al.*, *Science*, 247:1465-1468, 1990.

Wu *et al.*, *J. Biol. Chem.*, 263:14621-14624, 1988.

25 Wu *et al.*, *J. Biol. Chem.*, 264:16985-16987, 1989.

Yang *et al.*, *Arteriosclerosis*, 9(1):96-108, 1989.

Yang *et al.*, *Nature*, 323:738-742, 1986.

Yang, *Proc. Natl. Acad. Sci. USA*, 87:5523-5527, 1990.

Ye *et al.*, *J. Biol. Chem.*, 263(13):6337-6343, 1988.

30 Young *et al.*, *Arterioscl.*, 6:178-188, 1986.

Yu *et al.*, *Cell*, 76:933-945, 1994.